DFRWS USA 2025 - Selected Papers from the 25th Annual Digital Forensics Research Conference USA

# SoK: Timeline based event reconstruction for digital forensics: Terminology, methodology, and current challenges

Frank Breitinger [a,*], Hudan Studiawan [b], Chris Hargreaves [c]

[a] *Chair for Cybersecurity, University of Augsburg, Augsburg, Germany*
[b] *Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*
[c] *Department of Computer Science, University of Oxford, Oxford, United Kingdom*

## ARTICLE INFO

## ABSTRACT

Event reconstruction is a technique that examiners can use to attempt to infer past activities by analyzing digital artifacts. Despite its significance, the field suffers from fragmented research, with studies often focusing narrowly on aspects like timeline creation or tampering detection. This paper addresses the lack of a unified perspective by proposing a comprehensive framework for timeline-based event reconstruction, adapted from traditional forensic science models. We begin by harmonizing existing terminology and presenting a cohesive diagram that clarifies the relationships between key elements of the reconstruction process. Through a comprehensive literature survey, we classify and organize the main challenges, extending the discussion beyond common issues like data volume. Lastly, we highlight recent advancements and propose directions for future research, including specific research gaps. By providing a structured approach, key findings, and a clearer understanding of the underlying challenges, this work aims to strengthen the foundation of digital forensics.

## 1. Introduction

Event reconstruction involves recreating past events by analyzing digital artifacts, allowing examiners to determine system activities and make informed conclusions about what occurred. While traditional forensic science benefits from a well-defined framework summarizing the field (Ribaux, 2023), event reconstruction in digital forensics is often discussed in fragmented terms focusing on tasks such as super timeline creation (Gudhjónsson, 2010; Metz et al., 2024), tampering detection (Palmbach and Breitinger, 2020; Studiawan and Sohel, 2021) or environmental peculiarities (Schatz et al., 2006). As a result, research has centered on these narrow aspects, leaving broader challenges underexplored or overlooked. The absence of a unified perspective has led to a proliferation of terms, making it difficult to discuss event reconstruction comprehensively or find relevant research, e.g., some studies use the term artifact (Harichandran et al., 2016), others refer to observable facets (Jaquet-Chiffelle and Casey, 2021). Terms such as events (Carrier and Spafford, 2004a), user actions, interactions, or clicks (Neasbitt et al., 2014) are inconsistently used in literature.

*The three contributions:* First, the article discusses concepts and definitions in timeline-based event reconstruction and integrates them into a new visual model (the timeline-based event reconstruction model or *TER-Model*), divided into four quadrants, integrating digital forensic timeline-based terminology and Ribaux (2014)'s model. Second, with this delineation, we provide a thorough discussion of the issues associated with timeline-based event reconstruction. These issues can be used to evaluate event reconstructions and identify areas of uncertainty in the results. They can also be used to systematically identify weaknesses in the timeline generation and analysis techniques and contribute to a knowledge base of weaknesses such as SOLVE-IT (Hargreaves et al., 2025). Third, we provide future research directions needed within each quadrant of the event reconstruction process. This paper is predominantly theoretical, aiming to harmonize timeline-based event reconstruction terminology, however, a practical illustration of the use of the model is available online.[1]

*Not in scope:* The identification of relevant devices (computer profiling, Marrington et al. (2007)), legal constraints or ethical issues (Losavio et al., 2015), technical challenges such as encryption, sophistication of crime (Karie and Venter, 2015), or very general challenges, e. g., that "results must be reproducible and verifiable" (Soltani and Seno, 2019) are not in scope.

*Outline:* The next section summarizes core works in event

---

* Corresponding author.
*E-mail addresses:* frank.breitinger@uni-a.de (F. Breitinger), hudan@its.ac.id (H. Studiawan), christopher.hargreaves@cs.ox.ac.uk (C. Hargreaves).
[1] https://github.com/chrishargreaves/TER-model-example

reconstruction which serves as a foundation for this work. Subsequently, Sec. 3 presents terms and technology in existing literature and outlines the terminology used in this article. A contribution of this work is the TER-model which is developed and described in Sec. 4. Using the model, we identified challenges according to the methodology in Sec. 5 and organized the challenges for event reconstruction in the two main sections: Challenges stemming from environmental and process-related factors (Sec. 6) and Challenges stemming from deliberate interference (Sec. 7), which are summarized as key findings in Sec. 8. Considering these, Sec. 9 provides a discussion and identifies specific research gaps. The final section concludes the paper.

## 2. Event reconstruction

Lee et al. (2001) and many others have discussed event reconstruction for physical crime scenes. Carrier and Spafford (2004a,b) were the first to define it as applied in digital forensics and presented an event-based investigation framework. Their work defines the basic terminology and introduces a formal process model that mirrors physical crime scene investigations but is tailored to the unique aspects of digital evidence. We borrow from this work as discussed in Sec. 3.1.

Casey (2011)'s work includes the practicalities of linking evidence to behaviors and motives. Casey emphasizes three core analysis types: (1) temporal which helps establish the timeline of events (the focus of this article), (2) relational which explores the connections between objects, people, and locations, clarifying how different elements of the crime are related, and (3) functional which assesses what was possible or impossible, such as determining how a system or tool was used in the crime. Chabot et al. (2015a) defines terminology based on existing works, outlines challenges, and evaluates existing approaches. However, the authors limit their challenges to the volume of data and data heterogeneity where this article provides a broader discussion. Our work complements these existing works by providing a new visual model and a thorough discussion of challenges and future research.

## 3. Terminology

According to Neale (2023), there is a lack of harmonization in terms and definitions. This section briefly revisits (Sec. 3.1) and then highlights the terminology we use for this article (Sec. 3.2).

### 3.1. Terms and terminology in existing literature

Carrier and Spafford (2004a) define an event "as an occurrence that changes the state of one or more objects". Over time, researchers suggested to differentiate between low-level and high-level events (human-understandable) (Hargreaves and Patterson, 2012; Vanini et al., 2024) or introduced terms such as 'activity' (Marrington et al., 2007) or 'user–browser interaction' and 'click' which are used interchangeably by Neasbitt et al. (2014). Chabot et al. (2014) defines an event as "a single action occurring at a given time and lasting a certain duration".

Jaquet-Chiffelle and Casey (2021) define an event as "a complete collection of related things that have happened (or are happening) in a World within a specific closed interval of time. [...] The Event can be considered as a whole entity or as a collection of smaller sub-events". Notably, their framework emphasizes the role of traces and introduces several key concepts, including trace, facet, and observable facet. While these terms are well-established in forensic science (Ribaux, 2023), they are less common in digital forensics. Therefore, we adopt a different terminology, while drawing conceptual links to their work.

Similarly, the term *artifact* is used with different meanings. For instance, Harichandran et al. (2016) compares various definitions and concludes properties an artifact should have such as "artificiality/external force, antecedent temporal relation, and exceptionality". Horsman (2019) suggests "a digital object containing data which may describe the past, present or future use or function of a piece of software,

application or device for which it is attributable to". Casey et al. (2022) differentiates between atomic artifacts ("a singular unit of interpretable data that can be extracted from a given data source") and dependable artifacts ("one or more atomic artifacts needed to expose the atomic artifact of interest"). Lyle et al. (2022) extends the atomic artifact definition by adding "that is useful for addressing questions in forensic investigations", but assessing usefulness is difficult, subjective and may change over time.

### 3.2. Terminology used in this article

#### 3.2.1. Environments/systems

An environment/system is a computational setting or a software/hardware system that reacts to events such as user actions, API calls, or sensor inputs. Typically, it is one or more devices such as computers or smartphones but it could also be a virtual machine, network device, or cloud environment. For readability, the remainder of this paper uses the term environments instead of environments/systems. Note we use the plural, i.e., environments, considering that changes may be, locally, remotely, or both.

#### 3.2.2. Artifact

This article uses Casey et al. (2022) atomic artifact definition: a singular unit of interpretable data that can be extracted from a given data source. For simplicity, we will only say artifact throughout the paper. Examples include log files, registry keys, timestamps, or network traffic data.

#### 3.2.3. Event

Based on Jaquet-Chiffelle and Casey (2021), an event is "a complete collection of related things that have happened (or are happening) in a World within a specific closed interval of time." These can be treated as a singular entity or decomposed into smaller sub-events and cause environmental changes. This broad definition provides the flexibility for an event to be at the resolution of: 'file was accessed', 'Google search was performed', or 'user account was used to run a program' (consisting of at least two events: user logged in and user executed binary). Events can be triggered internally, e.g., a cron job, or externally such as someone clicking the mouse. Note that the distinction between event and sub-event is blurred and it is up to the user to define the granularity. For instance,

- an event is *sending an email* with sub-events such as opening the email client, typing, establishing a connection to the SMTP server, and sending the message, or
- an event is *establishing a connection to the SMTP server* with sub-events such as performing a DNS lookup, initiating a handshake, and authenticating the user credentials.

## 4. Model for event reconstruction

This work draws inspiration from Vanini et al. (2023), which, in turn, is influenced by the work of Ribaux (2023, p226, Fig. 4.4).[2] We adjusted these models to align with common digital forensics terminology and emphasize timeline-based event reconstruction (TER). Our model, named *TER-Model*, is depicted in Fig. 1 and can be separated into a *Reality* space (Sec. 4.2) and a *Reconstruction* space (Sec. 4.3). Each of these spaces can be further separated resulting in four quadrants (Q1-Q4). Before describing the model, this section first summarizes the goals of temporal event reconstruction which influenced the model. The summary of the systematization of knowledge (SoK) in the TER-Model is shown in Table 1.

---

[2] Note, this is an updated version from the previous work by Ribaux (2014) and thus has over a decade of history.
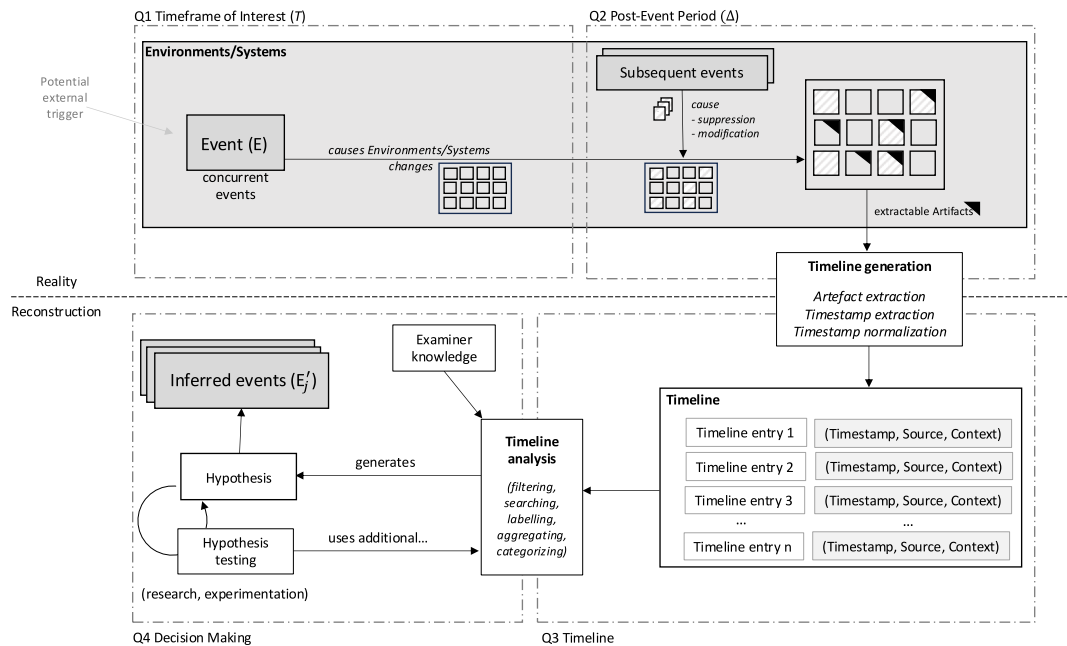
**Fig. 1.** TER-Model: Model of timeline-based event reconstruction in digital crime scenes. The small squares (3x4) in the upper part of the diagram represent changes caused by the primary event E (gray boxes) combined with additional changes from subsequent events (white–gray stripes).

### 4.1. Goals of temporal event reconstruction

Temporal event reconstruction aims to accurately recreate the sequence of events that occurred which includes finding gaps and inconsistencies, even if they cannot be accurately filled or corrected. Thus, it enables investigators to draw meaningful conclusions about what happened.

Event reconstruction involves several interrelated analytical processes that together provide a coherent and defensible narrative of what transpired. At its core it is temporal sequencing and correlation, where a precise order of events is created. It may be necessary to analyze their relationships across different timelines to uncover causal links, sequence dependencies, or concurrent activities (Adderley and Peterson, 2020). Beyond simple chronology, contextual analysis places these events within a broader framework, considering factors such as user behavior, system settings, or external influences to give the data deeper interpretive meaning (Chabot et al., 2015a). This groundwork supports hypothesis testing and scenario building, where investigators construct and refine possible explanations for what occurred, evaluating multiple narratives and ruling out those that conflict with the evidence (Willassen, 2008a, 2008b; Batten et al., 2012). It is crucial that the reconstructed timelines are confirmed through correlation and verification of evidence to ensure consistency and reliability. The goal is to produce a report to support legal proceedings that not only stands up to technical scrutiny, but also serves court proceedings by providing a clear, accurate and accessible story for stakeholders such as lawyers or jurors (Chabot et al., 2014; Xu and Xu, 2022).

### 4.2. Reality and its two dimensions (Q1, Q2)

#### 4.2.1. Q1: Timeframe of interest T

This quadrant is an interval that has a start time $t_S$ and an end time $t_E$, i.e., $T = [t_S, t_E]$ during which the event ($E$) and sub-events ($e_1, e_2, ...e_m$) occurred. Each $E$ or $e$ causes multiple environmental changes, e.g., new log entries, modified registry values, files marked as non-allocated, or updated timestamps.

The event ($E$) is what we wish to be able to say something about through the event reconstruction process. Carrier (2006) describes that

an event can be any "an occurrence that changes the state of the system" and Hargreaves (2009) continues that "digital events occur on a system often as a result of interactions with another digital device, or as a result of interactions with the real world". However, in Jaquet-Chiffelle and Casey (2021) event is formalized such that these external triggers are integrated into the event itself, defining an event that can capture the very broad, or the very detailed. In addition, there are *concurrent events* such as antivirus scanning files resulting in changes not tied to the primary event.

#### 4.2.2. Q2: Post-event period (Δ)

During this interval Δ, the environment-changes caused by $E$ may become intermingled with, altered, or overwritten by an ensemble of other data generated by unrelated *subsequent events*. Jaquet-Chiffelle and Casey (2021) categorized these changes as adjunction, suppression, and change. This second interval ends at time $t_P$ when the data is preserved/extracted, i.e., $\Delta = (t_E, t_P]$. As $t_E$ belongs to $T$, we exclude it here from this interval using a half-open interval. It is important to note that not all environment-changes can be extracted, such as missing/deleted files or new artifacts without a parser. These gaps may stem from many causes, for example a lack of knowledge in digital forensics, a tool setup, or errors in the timeline generation process. Hence, what can be extracted is named *extractable artifact*, which is therefore context specific.

#### 4.2.3. Timeline generation

Combined with preservation and acquisition, timeline generation bridges the *Reality* and *Reconstruction* spaces. Hargreaves et al. (2024b) define it as a process within a forensic analysis tool for "extracting timestamps from the file system … [and] applying file specific processing and extracting timestamps from within files such as the Windows Registry, log files, SQLite databases etc., that contain timestamps". This artifact and timestamp extraction is complemented by normalization, which is required since timestamps exist in a variety of formats (e.g., ASCII in a log vs. little-endian hexadecimal in a proprietary format), and resolutions (i.e., hours, minutes, seconds, nanoseconds, etc.) depending on their source (Raghavan and Saran, 2013). They may also be stored in UTC or local time. Ideally, after normalization, all timestamps should be

**Table 1**
Summary of systematization of knowledge (SoK) for timeline-based event reconstruction (TER).

| Paper | Focus area | Contribution type/ Challenge | TER quadrant | | | | Data source category | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Q1 | Q2 | Q3 | Q4 | Physical | File system | Multi sources | Logs | Other | Timestamp | Analysis | Mobile/ IoT | Volatile | Network |
| **Sec. 2 Event reconstruction** | | | | | | | | | | | | | | | | |
| Lee et al. (2001) | Foundational event reconstruction | Conceptual framework | ✓ | | | ✓ | • | | | | | | | | | |
| Carrier and Spafford (2004a,b) | Event-based investigation process | Process model | ✓ | ✓ | | | | | • | • | | | | | • | • |
| Casey (2011) | Temporal, relational analysis | Analytical framework | | | ✓ | ✓ | | | • | | | | | | | |
| Chabot et al. (2015a) | Terminology, data volume | State-of-the-art review | ✓ | | | ✓ | | | • | | | | | | | |
| Adderley and Peterson (2020) | Temporal sequencing | Timeline correlation | | | ✓ | | | | | • | | • | • | ○ | ○ | ○ |
| Willassen (2008a,b) | Hypothesis testing | Model-based reconstruction | | | | ✓ | | | | • | | • | | | | ○ |
| Batten et al. (2012) | Hypothesis development | Reasoning methodology | | | | ✓ | ○ | | | • | | ○ | | | | |
| Xu and Xu (2022) | Knowledge graph reasoning | Visualization and reasoning model | | | | ✓ | | | • | | | | | | | |
| **Sec. 3 Terminology** | | | | | | | | | | | | | | | | |
| Neale (2023) | Artifact terminology harmonization | Systematic terminology review | | ✓ | | | | | • | | | | | | | |
| Carrier and Spafford (2004a,b) | Event-based investigation process | Process model | ✓ | ✓ | | | | | • | • | | | | | • | • |
| Hargreaves and Patterson (2012) | Event granularity | Event granularity | | ✓ | ✓ | | | | • | • | | • | | ○ | | ○ |
| Marrington et al. (2007) | Computer activity | Activity terminology | | ✓ | ✓ | | | | • | • | | • | | | | |
| Neasbitt et al. (2014) | User interaction terminology | Interaction terminology | | | ✓ | | | | | ○ | | | | | | • |
| Chabot et al. (2014) | Duration-based event definition | Terminology refinement | ✓ | | | ✓ | | | • | | | | | | | |
| Jaquet-Chiffelle and Casey (2021) | Forensic event structure | Forensic event model | ✓ | | | ✓ | | | • | | | | | | | |
| Harichandran et al. (2016) | Artifact properties analysis | Artifact comparison | | ✓ | | | | | • | | | | | | | |
| Horsman (2019) | Artifact as digital object | Practical definition | | ✓ | | | | • | | • | | | | | | |
| Casey et al. (2022) | Artifact definition | Artifact catalog | | ✓ | | | | | • | | | | | | | |
| Lyle et al. (2022) | Artifact identification | Digital investigation techniques | | ✓ | | | | | • | | | | | | | |
| **Sec. 4 Model for event reconstruction** | | | | | | | | | | | | | | | | |
| Ribaux (2014, 2023) | Forensic trace model | Trace-based model | | ✓ | | | • | | | | | | | | | |
| Vanini et al. (2023) | Event source reliability | Reliability modeling | ✓ | | | ✓ | | | • | | | | | | | |
| Vanini et al. (2024) | Time anchor model | Timestamp interpretation framework | ✓ | | | ✓ | | | | • | | • | | | | ○ |
| Carrier (2006) | Investigation process model | Hypothesis-based model | | | | ✓ | • | | | ○ | | | | | | |
| Hargreaves (2009) | Evidence reliability testing | Reliability criteria | | | | ✓ | | | • | | | | | | | |
| Jaquet-Chiffelle and Casey (2021) | Event structure | Formal event model | ✓ | ✓ | | | | | • | | | | | | | |
| Hargreaves et al. (2024b) | Tool transparency | Tool capability model | | ✓ | ✓ | | | | • | | | | • | | | |

**Table 1** (*continued*)

| Paper | Focus area | Contribution type/Challenge | TER quadrant | | | | Data source category | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Q1 | Q2 | Q3 | Q4 | Physical | File system | Multi sources | Logs | Other | Timestamp | Analysis | Mobile/IoT | Volatile | Network |
| Raghavan and Saran (2013) | Timestamp interpretation | Timestamp model | ✓ | | ✓ | | | | | • | | • | | ○ | | • |
| Hargreaves and Patterson (2012) | Timeline generation model | Timeline generation model | | ✓ | ✓ | | | • | | • | | • | | ○ | | ○ |
| Studiawan et al. (2020a); Studiawan (2023) | Event abstraction | Event abstraction model | | | ✓ | ✓ | | | | • | | • | • | ○ | | ○ |
| Carrier and Spafford (2004a,b) | Hypothesis-based investigation | Hypothesis model | ✓ | | | ✓ | • | | | | | | | | | |
| Gladyshev and Patel (2004) | Event inference | FSM reconstruction | | | | ✓ | | | | • | ○ | | | | | ○ |
| Amato et al. (2017) | Semantic evidence correlation | Ontology-based model | | ✓ | ✓ | | | | | • | | | | | | |
| Xu and Xu (2022) | Knowledge graph presentation | Reasoning model | | | ✓ | ✓ | | | • | | | | | | | |
| **Sec. 6 Challenges stemming from environmental and process-related factors** | | | | | | | | | | | | | | | | |
| **Sec. 6.1.1 Incorrect environment time** | | | | | | | | | | | | | | | | |
| Stevens (2004) | Misconfigured system clocks | Clock drift challenge | ✓ | | | | | | | | | • | | | | |
| Raghavan and Saran (2013) | Timestamp normalization and storage issues | Timestamp interpretation framework | ✓ | | | | | | | • | | • | • | ○ | • | • |
| Vanini et al. (2024) | Time anchor abstraction model | Time anchor modeling | ✓ | | | ✓ | | | | • | | • | | | | ○ |
| Kaart and Laraghy (2014) | Incorrect timezone data handling | Time zone configuration | ✓ | | | | | | | • | | • | • | • | • | • |
| Schatz et al. (2006); Buchholz and Tjaden (2007) | Network-induced skew, unsync clocks | Distributed system time consistency | ✓ | | | | | | | • | | • | • | ○ | | • |
| Henderson (2009) | Clock skew in shared environments | Network delay and skew | ✓ | | | | | | | • | | • | | ○ | | • |
| **Sec. 6.1.2 Configurations and implementations** | | | | | | | | | | | | | | | | |
| Adedayo and Olivier (2015) | Log suppression, redirection | Log misconfiguration | ✓ | | | | | • | | • | | ○ | | | | |
| Fernández-Fuentes et al. (2022) | Absence of traceability in apps | Limited logging capability | ✓ | | | | | | | ○ | ○ | | | | • | • |
| **Sec. 6.1.3 Environmental anomalies** | | | | | | | | | | | | | | | | |
| Studiawan et al. (2019) | Unrecoverable system restarts | Environmental disruption | ✓ | | | | | | | • | | • | • | ○ | ○ | ○ |
| Oh et al. (2022) | Sudden device restarts | Restart-induced log gaps | ✓ | | | | | • | | • | | • | | ○ | | |
| Marrington et al. (2011) | Program faults, data corruption | Software instability | ✓ | | | | | • | | • | | • | • | | | |
| **Sec. 6.1.4 Data fluctuation** | | | | | | | | | | | | | | | | |
| Sandvik et al. (2021) | Short lifespan of traces | Volatile trace loss | ✓ | | | | | | | | | | | | • | |
| Marangos et al. (2016) | Evidence affected by operational cycles | Temporal instability | ✓ | | | | | | | | | | | | • | |
| **Sec. 6.2 Post-event period** | | | | | | | | | | | | | | | | |
| Gruber et al. (2023) | Evidence altered during acquisition | Contamination challenge | | ✓ | | | | • | | • | | • | | | • | • |
| Jaquet-Chiffelle and Casey (2021) | Evidence fragility and impermanence | Temporal evidence integrity | | ✓ | | | | • | | | | | | | | |
| Khan et al. (2007) | Overwriting of data, log aging | Aging challenge | | ✓ | | | | • | | • | | • | | | ○ | |

**Table 1** (*continued*)

| Paper | Focus area | Contribution type/Challenge | TER quadrant | | | | Data source category | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Q1 | Q2 | Q3 | Q4 | Physical | File system | Multi sources | Logs | Other | Timestamp | Analysis | Mobile/IoT | Volatile | Network |
| Soltani et al. (2019); Schuster (2007) | Metadata decay, inaccuracy | Artifact degradation | ✓ | | | | | • | • | • | | • | | | ○ | ○ |
| **Sec. 6.3 Timeline** | | | | | | | | | | | | | | | | |
| Patterson and Hargreaves (2012) | Cross-source correlation | Source integration challenge | | ✓ | | | | • | • | • | | • | • | | | |
| Mohammed et al. (2016) | Data format diversity | Data normalization challenge | | ✓ | | | | • | • | • | | • | • | • | • | • |
| Horsman (2019) | Artifact parsing complexity | Parser dependency challenge | | ✓ | | | | • | | • | | | | | | |
| Soltani and Seno (2017) | Missing/incomplete timestamps | Extraction incompleteness | | ✓ | | | | | • | | | | | | | |
| Gómez et al. (2005); Levett et al. (2010) | Correlation of heterogeneous data | Multi-source correlation | | ✓ | | | | | • | | | | | | | |
| Kälber et al. (2013); Hargreaves et al. (2024b) | Tool transparency and automation limitations | Human-tool balance challenge | | ✓ | ✓ | | | | • | | | | | | | |
| Bhat et al. (2021) | Misconfigured analysis environments | Tool setup challenge | | ✓ | ✓ | | | | • | | | | • | | | |
| **Sec. 6.4 Decision making** | | | | | | | | | | | | | | | | |
| Chabot et al. (2015a) | Data volume for timeline analysis | Scalability and overload challenge | | | | ✓ | | | • | • | | | | | | |
| Quick and Choo (2014) | Computational resource limitations | Resource requirement challenge | | | | ✓ | | | • | • | | | | | | |
| Buchholz and Falk (2005) | Event aggregation | Event abstraction for analysis | | | | ✓ | | | • | • | | | | | | |
| Kiernan and Terzi (2009) | Event summarization | Abstraction and streamlining | | | | ✓ | | | | • | | | | | | |
| Osborne and Turnbull (2009) | Visualization accuracy | Visual representation integrity | | | | ✓ | | | | | | | • | | | |
| **Sec. 7 Challenges stemming from deliberate interference** | | | | | | | | | | | | | | | | |
| Casey (2020) | Strength and scale of inference | Evaluative opinion framework | | | | ✓ | | | | | • | | • | | | |
| Vanini et al. (2024) | Time manipulation, clock tampering | Timeframe manipulation | ✓ | | | | | | | • | • | | | | | ○ |
| MITRE (2023) | Environment manipulation, disabled logging | Environment tampering | ✓ | | | | | | | • | | | | | | |
| Conlan et al. (2016) | Erasure or alteration of evidence using tools | Anti-forensics tool usage | ✓ | | | | | | • | | | | | | | |
| Palmbach and Breitinger (2020) | File and log manipulation using malware | Malware-assisted anti-forensics | ✓ | | | | | | • | | | | | | | |
| Malhotra et al. (2015) | Service manipulation (e.g., NTP tampering) | Service compromise | ✓ | | | | | | | | | | | | | • |
| Choi et al. (2021) | Post-event manipulation: logs, timestamps, files | Artifact modification & deletion | | ✓ | | | | | | • | | | | | | |

Notes: • Mentioned in the paper ○ Not specifically mentioned, but can be implemented using the data source.

presented in the same format for better readability and sortability.

### 4.3. Perception

The lower section of the diagram represents how examiners attempt to reconstruct past events using reasoning and available evidence. This process involves uncertainty, as the past cannot be revisited, making absolute certainty unattainable.

#### 4.3.1. Q3: Timeline

Examiners construct a timeline to facilitate analysis, and the DFPulse 2024 Practitioner Survey (Hargreaves et al., 2024a) reports 80.3 % are using timelines 'often' or 'almost always'. Timelines are composed of a series of entries, each derived from individual artifacts that are arranged chronologically. Artifacts may originate from multiple independent data sources, e.g., a computer and a smartwatch. While specific implementations store multiple data points per event, fundamentally these *timeline entries* are defined as a 3-tuple ($t$, $S$, $C$):

- The normalized timestamps ($t$) are used to order the timeline chronologically.
- A source $S$ refers to the specific location from which the timestamp and context originate, such as the Master File Table (MFT), Windows registry, EPROCESS block in memory, or Chrome browser history file. For clarity, $S$ should be as detailed as possible; instead of stating the registry, the exact registry key path should be specified.
- A context $C$ defines what the timestamp represents, such as the modification timestamp within the Standard Information Attribute (SIA) of MFT entry, or a value in a specific row or field within a database. Given the wide variety of contexts, a generic term is used to encompass the diverse nature of these representations.

These timeline entries should not be conflated with events themselves or 'low-level events' (Hargreaves and Patterson, 2012). The context provided by each entry, such as a value in a 'modified' or 'last change' field within a file system structure, does not inherently represent a specific event, such as a file modification. Instead, it reflects environmental behavior that must be understood before making any assumptions about what event occurred. This distinction is critical: while timeline entries provide the raw data needed for event reconstruction, they are not events in and of themselves. Rather, they are normalized, sorted compilations of data that result from parsing artifacts left by events. Therefore, we argue that the term event should be reserved for the inferred actions, while the term timeline entry more accurately describes the data points that examiners use to reach those inferences.

#### 4.3.2. Timeline analysis

Timeline analysis bridges Q3 and Q4, and describes the process of moving from having a timeline to reconstructing events, which uses refinement techniques such as: filtering irrelevant entries, highlighting key entries, or aggregating entries into more meaningful events (Hargreaves and Patterson, 2012). Several other concepts have been discussed such as event abstraction (Studiawan et al., 2020a; Studiawan, 2023), the application of machine learning (Khan and Wakeman, 2006), or visualization (Berggren et al., 2024; Debinski et al., 2019). Timeline analysis also draws in *examiner knowledge* to understand potential events that are capable of producing the timeline entries and integrating them into a reasoning process (Gladyshev and Patel, 2004).

#### 4.3.3. Q4: Hypotheses and event inference

To accurately approach event reconstruction, it is essential to distinguish between the event $E$ that occurred in reality and the inferred event $E'$ which is derived from the analysis of timeline entries. In the context of hypothesis generation, $E'$ represents the best approximation based on the available evidence. We define an inferred event $E'$ as a *reconstructed scenario that may have occurred within a specific time frame, based on the interpretation and analysis of timeline entries and associated artifacts.* This definition acknowledges the uncertainty in reconstructing past events.

Consideration of the timeline entries in the context of examiner knowledge may result in multiple plausible scenarios (Jaquet-Chiffelle and Casey, 2021; Gladyshev and Patel, 2004). Hargreaves (2009) states "if there are multiple events that could cause the same state of digital data, there is an actual, true event that caused it, and one or more other events that did not." This means that rather than arriving at a single definitive inferred event $E'$, we may generate $k$ alternative events, denoted as $E'_j$ where $1 \leq j \leq k$. Each $E'_j$ represents a distinct interpretation of the evidence, each of which could potentially explain the observed data. These multiple instances of $E'$ highlight the complexity and ambiguity, where different sequences of events could produce similar artifacts. The process involves not only constructing these alternatives but also systematically and repeatedly testing and eliminating hypotheses to converge on the most likely scenario while acknowledging that multiple interpretations may still be viable based on the available evidence. To test and eliminate hypotheses, Casey (2020)'s 'Strength of evidence scale' (C-Scale) may be used, and it may involve research into artifact interpretation and experiments to determine if a set of actions could produce the observed system changes.

### 5. Methodology for challenge identification

To identify and categorize the challenges in event reconstruction, we followed a structured literature review process designed to balance breadth with relevance. The goal was not to exhaustively capture all existing work but to obtain a representative and insightful overview of the key challenges discussed in the field.

**Search strategy:** We defined a set of core search terms related to the topic: event reconstruction, timeline, timestamp analysis, digital forensics, correlation, challenges, and problems. These terms were combined using Boolean operators and phrasing variations (e.g., quotation marks for exact matches). Searches were conducted using Google Scholar, which indexes most major academic publishers (e.g., IEEE, ACM, Wiley, Springer) and relevant platforms such as DFRWS. org and arXiv.

**Selection criteria:** For each query, we considered the first two pages of results (i.e., 20 entries). Articles were initially screened based on metadata displayed: title, author(s), publication venue, and two-line extract. If no direct reference to digital forensics was evident, the article was discarded. This filtering yielded a preliminary pool of approximately 200 articles.

**Challenge extraction:** We extracted mentions of challenges primarily from the abstract and introduction sections, where such content is frequently summarized. Targeted keyword searches (e.g., challenge, problem, limitation) were also used within full texts to uncover implicit references.

**Classification:** The identified challenges were then mapped onto a diagram, categorizing them according to the stage or context in which they occur within the event reconstruction process.

We also incorporated our domain expertise to address gaps in the literature, recognizing that some relevant challenges may not have been explicitly highlighted in existing works.

### 5.1. Limitations

The article collection and analysis were conducted manually, which may have led to the omission or misclassification of relevant articles. By restricting searches to Google Scholar and considering only the first two pages of results, important sources further down the list or from other

databases may have been excluded. The focus on abstracts and introductions might have caused us to overlook challenges discussed deeper within the papers. Moreover, the subjective nature of challenge classification introduces potential bias based on the researchers' interpretations. Finally, the absence of automated or statistical tools for extraction and categorization limits the objectivity and comprehensiveness of the analysis. Despite these limitations, we believe the following sections offer a comprehensive and nuanced overview of the challenges.

## 6. Challenges stemming from environmental and process-related factors

This section focuses on *unintentional* challenges and the structure follows the diagram's flow, discussing each quadrant.

Note, that while we have strived to define the challenge categories as distinctly as possible, some overlap is inevitable due to the interconnected nature of these activities. Certain actions may reasonably fall into multiple categories, depending on the context. The categorization is designed to provide guidance rather than enforce strict mutual exclusivity.

### 6.1. Q1: Timeframe of interest

Four areas have been identified:

#### 6.1.1. Incorrect environment time

Clock-related challenges originate from the system time which is used to derive timestamps. If the clock is incorrect, all timestamps originating from this clock are incorrect (Stevens, 2004; Raghavan and Saran, 2013; Vanini et al., 2024).

**Clock skew:** Skew refers to the difference in time readings between different systems. One reason for clock skew could be propagation delays which may occur due to network delays (Schatz et al., 2006; Henderson, 2009) or due to synchronization problems, e.g., NTP servers providing incorrect times (Buchholz and Tjaden, 2007; Hampton and Baig, 2016).

**Clock drift:** Drift is the gradual deviation of a clock from the correct time, often caused by factors such as changes in temperature, voltage fluctuations, or inherent defects in the clock circuitry (Sandvik and Årnes, 2018). Clock drift may exacerbate over time. As drift accumulates, the discrepancies between different systems' clocks can grow, making it increasingly difficult to correlate events across environments (Becker et al., 2008).

**Time zone changes:** As systems traverse different time zones, whether due to travel or daylight-saving time changes, the system time may change (Stevens, 2004). This adjustment process can also be error-prone, e.g., due to an inaccurate time zone database (Kaart and Laraghy, 2014). Compared to skew and drift, the range is significantly larger, i.e., hours instead of seconds. Typically this is only relevant where local time is stored in a data structure rather than storing UTC.

Note that virtual environments come with their challenges which are beyond the scope of this article but have been discussed in VMware (2008).

#### 6.1.2. Configurations and implementations

Environments, systems, and application configurations define how/what data is generated, stored, and logged. These configurations comprise a wide range of settings, including logging levels, storage policies, network settings, and security controls.

**Suppression/deletion:** Conservative default settings can result in insufficient logging, leading to missing artifacts, e.g., database logs prioritizing space efficiency over detail (Adedayo and Olivier, 2015). Systems may also be configured to suppress artifacts, such as private browsing (Fernández-Fuentes et al., 2022), or delete them, such as printer jobs removed after completion (Gladyshev and Patel, 2004) or

when an application is closed.

**Inconsistent implementations:** Different resolutions lead to inconsistencies, e.g., timestamps recorded in hh:mm vs. hh:mm:ss format (Song et al., 2016). File systems, drivers, and implementations may behave differently leading to unpredictable behavior (Bang et al., 2009; Nordvik and Axelsson, 2022).

#### 6.1.3. Environmental anomalies

Environments may not behave as expected leading to the destructing of evidence or the not-creation of artifacts:

**(OS) Crashes:** A crash (system, application) can result in the loss or corruption of artifacts, potentially leaving logs incomplete and missing key events (Studiawan et al., 2019; Oh et al., 2022). Detecting crashes can be challenging, particularly if the logging mechanisms themselves are compromised during the crash. Crashes may also lead to restart anomalies such as services or applications that are supposed to start automatically failing to do so potentially altering the way subsequent events are logged.

**Software bugs:** Bugs in software may cause errors in data logging, such as incorrect timestamps or missing events (Marrington et al., 2011).

**Resource exhaustion and failure:** Environments under heavy load may fail to log events properly due to resource constraints, leading to delayed or missed entries in the event data. Failures, including hardware malfunctions, can lead to inadequate data (Marrington et al., 2011).

#### 6.1.4. Data fluctuation

Data may not be accessible due to or only with additional burden:

**Data volatility:** Volatile data, such as RAM content or network traffic, is lost if the Δ is too large. In addition, IoT devices often have resource constraints resulting in short-lived data (Sandvik et al., 2021). In cloud environments, VMs can be easily deleted including their logs (Marangos et al., 2016).

**Environment bounds:** The changes resulting from an event may be distributed across multiple locations, including cloud environments, resulting in fragmented evidence that is challenging to collect and analyze (Herman et al., 2020; Joseph and Singh, 2019; Manral et al., 2019).

Even with the cooperation of external service providers, data cannot be recovered, particularly when logging is explicitly disabled, as is often the case with many VPN services.

### 6.2. Q2: Post-event Period

This period relates to the influence of time on the changes left behind after an event.

#### 6.2.1. Subsequent events impacting changes

Over time the changes generated by the primary event are altered by subsequent events (referred to as intrinsic events by Jaquet-Chiffelle and Casey (2021), or evidence dynamics by Gruber et al. (2023)).

**Deletion:** Initial changes may disappear due to subsequent events. Examples are rotating logs (Sandvik et al., 2021), temporary files, routine cleanup tasks, or reboots.

**Alteration/overwriting:** Subsequent events can modify or replace existing data. For instance, Khan et al. (2007) mention that much of the application footprint is rewritten each time the application runs. Routine file operations, such as automatic backups or updates, may also overwrite metadata, configurations, or timestamps (Soltani et al., 2019).

#### 6.2.2. Aging and degradation

Digital artifacts and physical devices are susceptible to degradation, affecting their reliability and accessibility. This degradation can manifest as file corruption, obsolescence of file types, or the deterioration of storage media. Furthermore, changes in software, file formats, or logging systems can introduce additional challenges. As schemas evolve, inconsistencies in log formats may emerge, complicating the process of

reconciling older and newer data entries. Backward compatibility issues also arise when outdated systems or logs are incompatible with modern tools, requiring extra effort to ensure that historical data remains interpretable and consistent across different versions (Schuster, 2007).

### 6.3. Q3: Timeline

This third quadrant summarizes all timeline-related challenges. We decided to include the trans-boundary boxes, i.e., timeline generation (Q2-Q3) and timeline analysis (Q3-Q4), in this section as we think they are closer related to the timeline.

#### 6.3.1. Timeline generation

Data comes from various systems, including traditional computing environments and a growing number of IoT devices, each with distinct structures, conventions, and formats (Patterson and Hargreaves, 2012; Mohammed et al., 2016). This increasing *heterogeneity* of both data sources and devices causes several challenges.

**Artifact/timestamp extraction:** Extracting data presents an ongoing challenge, as tools must be continuously updated to accommodate new and evolving software (Horsman, 2019). The acquisition process can introduce alterations, particularly when conducted on live systems, such as during memory dumps (Soltani and Seno, 2017; Gruber et al., 2023).

**Normalization:** This involves converting diverse data types, such as logs, databases, and sensor outputs, into a standardized structure that enables comprehensive analysis (Han et al., 2020). This can be challenging due to different timestamp formats, timestamp resolutions, and timezone settings. Timestamp formats can also change over time, meaning timestamp normalization needs to be updated over time and handle older and newer formats.

**Contamination and process problems:** Evidence might be unintentionally modified during collection or handling, e.g., failing to use a write blocker (Gruber et al., 2023) or corrupt software, leading to data contamination. Similarly, lapses in maintaining a proper chain of custody can result in evidence being mishandled, misplaced, or questioned in terms of authenticity and reliability.

**Source combination:** Combining data from multiple sources to create a unified perception is challenging, especially when sources have different levels of reliability or granularity (Gómez et al., 2005; Levett et al., 2010).

#### 6.3.2. Tool capabilities and usage

Balancing automated tools with manual analysis is essential yet challenging. While automation expedites the process, it may overlook nuances that a human analyst would catch (Kälber et al., 2013) and can introduce various types of error (Hargreaves et al., 2024b).

**Usage challenges:** Incorrect settings or carelessness can lead to incorrect results. For example, errors in the configuration of the tools have been shown to result in inaccurate extractions of digital evidence, which can impact the credibility of the findings (Bhat et al., 2021). The transition to a new tool may lead to misinterpretation as tools may interpret/visualize data differently. Some features of tools also do not help in reducing chances of investigator misinterpretation (see Hargreaves et al. (2024b)), e.g., if a tool provides an automated result of a Google search occurring, this is easy to interpret the event occurring as a fact rather than Google search data being present. This is an event reconstruction process, with all the uncertainty that could be present, as discussed in Sec. 6.4. Tools can conflate facts with interpretation within their interfaces.

**Transparency:** Many tools operate as black-boxes making it unclear how artifacts are handled. Transparency of functionality is critical, as proprietary processes can influence assumptions or conclusions, leading to misinterpretation.

**Handling volume:** Tools may have limits on the amount of data they can process or the complexity of queries, leading to unnoticed gaps in analysis, e.g., a tool limited to analyzing 5000 files at once. Consequently, validation is essential, but challenging, given the rapid change of artifacts (Horsman, 2018; Arshad et al., 2018).

**AI-powered examination:** AI-powered tools introduce complexities regarding explainability and transparency, not just of the models but of training data. Recent approaches such as LLMs are also problematic due to their non-deterministic nature and in many cases opaque training data and processes. These tools can produce inaccurate or misleading outputs, such as AI-generated errors or 'hallucinations' which can affect the analysis (Scanlon et al., 2023).

Developers aiming to create tools should consider the seven criteria outlined by Chabot et al. (2015b), which provide a comprehensive framework for ensuring an efficient reconstruction tool.

### 6.4. Q4: Decision making

Q4 involves the generation and testing of hypotheses based on the timeline. This is critical and Hargreaves (2009) goes as far as defining a digital investigation as "a process that formulates and tests hypothesis using digital evidence" with the prior stages facilitating this goal. Some areas of this are explored, e.g., timeline analysis, but others, such as hypothesis forming and testing are less frequently discussed.

#### 6.4.1. Timeline analysis

Although the processing is mostly done using tools, this section highlights challenges originating from the processing of timeline entries.

**Volume of data:** The extensive amount of information (number of entries in the timeline) makes the analysis time-consuming (Chabot et al., 2015a) and overloads examiners. Additionally, significant resources are needed to extract, process, and store this data, including computational power, storage capacity, and advanced data management tools (Quick and Choo, 2014).

**Aggregation, organization and visualization:** Techniques such as combining related events into cohesive units (sometimes called high-level events or super events) (Buchholz and Falk, 2005; Kiernan and Terzi, 2009; Hargreaves and Patterson, 2012; Inglot and Liu, 2014; Raju et al., 2017) can streamline analysis but may result in the loss of granularity or context. Similarly, visualizations (Osborne and Turnbull, 2009) require consideration to ensure that they accurately represent the data without oversimplifying or distorting the information. The volume of the raw data can be a challenge to visualize and reduction of the data before visualization is meaningful may be necessary, e.g., Hargreaves and Patterson (2012).

**Correlation:** The process of establishing meaningful relationships between disparate timelines entries is fraught with difficulties, especially when data originates from various sources or formats (Schatz et al., 2006) or times across environments are not synchronized (Marangos et al., 2016). Detecting and validating these connections requires experience and meticulous attention (Amato et al., 2017). For example, incorrect handling of local time vs. UTC can disrupt the sequencing of events, particularly in global systems where data spans multiple time zones (Buchholz and Tjaden, 2007). Verifying data across different sources and formats is challenging but necessary to ensure the accuracy and completeness of the reconstructed timeline.

#### 6.4.2. Interpretation, trust and integrity

Ensuring that data is accurate and trustworthy is fundamental (Neale et al., 2022). Determining which sources to trust and how to weigh them can significantly affect the reliability of the reconstruction. This challenge becomes even more pronounced when different sources report the same event but provide inconsistent or conflicting details, leading to uncertainty.

**Interpretation:** Investigators work with a static set of data which includes evidence and irrelevant information generated by subsequent activities or during investigative processes (Roux et al., 2022). Misinterpretation can arise from factors such as incorrect ordering,

aggregation, or filtering of entries, leading to distortions in the reconstructed narrative but also from unawareness of an examiner, i.e., insufficient knowledge of an event or timestamp (Boyd and Forster, 2004).

**Untrusted internal sources:** The presence of anti-forensic tools (Conlan et al., 2016) or tampering indicators, such as manipulated timestamps or hidden data, raises suspicion about the authenticity of the evidence.[3] According to Neale (2023), detecting and addressing such tampering is crucial to maintaining trust in the evidence (more in Sec. 7).

**Untrusted external sources:** Combining data from external sources, such as cloud services, introduces additional challenges. When the integrity of these sources cannot be independently verified, especially due to possible alterations in transit or at rest, the reliability of the event reconstruction may be compromised (Battistoni et al., 2016).

### 6.4.3. Knowledge and perception bias

Investigators may interpret evidence differently based on their prior knowledge, experience, or expectations, which can lead to skewed interpretations of the data. Perception and decision bias may cause certain patterns or details to be overlooked.

**Artifact interpretation knowledge:** Previous knowledge may become outdated due to the release of a new operating system, or new version of an application (Horsman, 2019). Examiners may be unaware of certain behaviors, e.g., Thierry and Müller (2022) identified multiple unexpected and non-compliant behaviors of timestamps. Limitations in knowledge reduce the investigator's ability to generate viable alternative hypotheses that would produce the same artifacts.

**Algorithmic bias:** Tools operate based on algorithms that might make certain assumptions or prioritize specific types of data, which can introduce biases into the reconstructed events (Jinad et al., 2024). For instance, an AI-powered tool may be biased due to unbalanced training data.

**Human bias:** Analysts may bring their own preconceptions into the analysis, influencing how they interpret and prioritize different events (Kang et al., 2013). This can lead to confirmation bias, where analysts might favor hypotheses that align with their pre-existing beliefs or expectations, unintentionally skewing the analysis (Kassin et al., 2013).

### 6.4.4. Complexity in testing hypotheses

Testing hypotheses against a timeline is complex, especially when considering all the aforementioned challenges.

**Multiple interpretations:** Evidence may be open to multiple interpretations, making it difficult to draw definitive conclusions and infer events from the past. This ambiguity can lead to varied interpretations of the same data, which impacts the ability to test hypotheses with certainty. Effective hypothesis testing must address temporal inaccuracies or manage the inherent uncertainty that arises from imperfect data such as log files (Latzo and Freiling, 2019).

**Defining error:** Hargreaves (2009) discusses that error in event reconstruction can be defined as "the difference between the inferred history and the true history of the examined digital evidence". This error cannot necessarily be expressed as a definite value, e.g., $x \pm y$, but can be expressed as uncertainty (possible error) in the inferred events, i.e., alternative possible hypothesized events that explain the current state of the examined digital evidence. Communicating these uncertainties transparently is vital to ensure that conclusions drawn are appropriately qualified and reflect the limitations of the available evidence.

## 7. Challenges stemming from deliberate interference

To complement the previous section, this one outlines challenges stemming from deliberate actions such as backdating, erasing, or wiping, to hide activities (Casey, 2020). While it may not always be the case, for this work we assume that the investigative body and tool vendors are free from insider threats. Therefore, challenges are limited to the *Reality space*.

As already pointed out in Sec. 6, some overlap of challenges is inevitable due to the interconnected nature of these activities.

### 7.1. Q1: Timeframe of interest

Interference with the environment can be conducted before the event occurs, with the intent to complicate investigations. Such interference often seeks to generate misleading artifacts or prevent their creation altogether, e.g., examples under 'defence evasion' in the MITRE ATT&CK Matrix[4].

**Time manipulation:** An adversary may turn off 'set time and date automatically' and actively manipulate the system time or timezone (Vanini et al., 2024). Even when detected, distinguishing between accidental misconfigurations and deliberate tampering remains difficult.

**Environment manipulation:** It is possible to disable or tamper with logging mechanisms, preventing activities from being recorded. Similarly, security tools may be compromised or altered (MITRE, 2023). Decoys such as fake accounts or planted traps such as cleanup scripts may be used to further obscure activities.

**Anti-forensics and malware:** Adversaries may use software to obscure their actions. For instance, anti-forensic tools erase or alter evidence (Conlan et al., 2016) or rootkits and malware to cover access and manipulations to files and logs (Palmbach and Breitinger, 2020). Anonymization services such as VPNs and TOR hide the attacker's origin, making it difficult to trace activities.

**Service manipulation:** Instead of manipulating an environment directly, an adversary may compromise utilized services. For instance, by manipulating the NTP service, an attacker can change the system time (Malhotra et al., 2015). Another example would be a compromised update server.

### 7.2. Q2: Post-event Period

In the post-event period, one may **manipulate or delete metadata or content** such as altering timestamps, modifying log entries, or deleting critical files (e.g., remote wiping of mobile devices). Logs and other files are often not protected against alternation or deletion (Choi et al., 2021). Active tampering and manipulation of artifacts present some of the most challenging obstacles in event reconstruction and the risk of misinterpretation increases (Casey, 2020) especially when performed from advanced persistent threads.

## 8. Key findings

This section summarizes the key findings identified in the foundational sections 2 to 4, and the challenge identification sections 6 and 7:

1. The terms "event" and "artifact" in digital forensics are defined inconsistently across existing studies and it leads to ambiguity in their usage.
2. Event reconstruction relies on modeling two critical intervals: the timeframe of interest ($T$) where events occur, and the post-event period ($\Delta$) where subsequent changes may overwrite or obscure evidence.

---

[3] We decided to include this challenge here and not in Sec. 7 (deliberate interference) as the presence of these tools does not necessarily mean that they were executed.

---

[4] https://attack.mitre.org/tactics/TA0005/

3. Event reconstruction is highly affected by unintentional challenges such as incorrect system time, insufficient logging, environmental anomalies, and data volatility.
4. Subsequent events can delete, overwrite, or degrade digital artifacts; so they reduce the availability and reliability of evidence over time.
5. Timeline generation faces challenges from data heterogeneity, software updates, extraction errors, normalization issues, and tool limitations.
6. Event reconstruction requires careful hypothesis generation and testing, but faces challenges from data volume, correlation complexity, trust issues, and investigator bias.
7. Deliberate actions such as time manipulation, anti-forensics, and post-event tampering can alter or destroy digital evidence and make event reconstruction even more challenging.
8. Several research directions have emerged to address challenges in event reconstruction, including forensic readiness, improved artifact extraction, timeline verification, tamper detection, AI/NLP integration, and advanced analysis techniques.

## 9. Discussion and research gaps

From the previous sections, the summary of key findings, and Table 1 (which provides a mapping of the focus areas in Sections 2 to 7, against the quadrants in Fig. 1, illustrating the distribution of existing research) it is possible to infer general research gaps. However, this section highlights selected significant challenges and proposes specific potential avenues for future research.

The section is organized by quadrant of the TER-model, demonstrating the utility of the model as an organizational tool. Given the vast body of literature, it is not feasible to reference every relevant article. Therefore, we focus on studies from our initial collection as well as recent works.

One general point is that throughout the TER-model (Q1-Q4) a broad research gap is the understanding and handling of uncertainty, from system configuration through to a reliance on examiner knowledge for hypothesis generation and testing. This is considered an ongoing limitation to the process that requires addressing.

> **Research Gap 1**. *Uncertainty is potentially introduced throughout the model and research into handling it at each stage, and how it could propagate is needed.*

### 9.1. Q1: Timeframe of interest

Digital forensic readiness is a proactive approach ensuring systems and networks are prepared to efficiently collect, preserve, and analyze evidence when a security incident occurs (Sachowski, 2019). Forensic readiness for event logging has been researched, as demonstrated by Reddy and Venter (2013) and Kebande and Venter (2018). To support forensic readiness, administrators should activate extended logging, which records additional data and audit trails. Moreover, operating system developers could still provide more comprehensive system-related logs (Rivera-Ortiz and Pasquale, 2019) but this conflicts with privacy-centric approaches expected from consumers.

This also has anti-forensics implications. If an attacker deletes logs (one of the primary sources for event reconstruction), investigators must first recover them (as discussed in Q2/Q3). To address this, security measures such as centralized or encrypted log servers could be implemented in systems where this is feasible, and even advanced techniques such as blockchain can be used to mitigate anti-forensic techniques (Kłos and El Fray, 2020).

> **Research Gap 2**. *Forensic readiness needs further development, and more creative solutions need researching to achieve similar goals on 'unmanaged' systems where forensic readiness solutions cannot be deployed.*

### 9.2. Q2: Post-event period

In evidence seizure, timing has an effect during forensic investigations. This affects if volatile artifacts are captured if not done on time, e.g., credentials stored in memory. Secondly, challenges related to cloud environments imply any delays in data acquisition may effortlessly cause the loss of crucial evidence, e.g., Alqahtany et al. (2016) discuss evidence that supports the need for timely acquisition. There is also the issue of long-term log retention by internet service providers, which may be important in some cases (Khan et al., 2016). Mandating extended retention ensures information can be accessed after an incident, but conflicts with privacy regulations. There are also 'awareness' concerns. For victim systems, communication is crucial to ensure device owners minimize interactions with devices containing potential evidence. The same applies to examiners, where changes to the evidence should be anticipated and minimized from a data preservation/acquisition perspective (Gruber et al., 2023). Moreover, recent work by Spichiger and Adelstein (2025) highlights that preservation should not be narrowly focused on the trace itself but must also consider the reference environment in which the trace was produced. As systems evolve, e.g., through software updates, operating systems, or third-party services, insufficient preservation of reference data can result in a loss of contextual meaning and increase the uncertainty of later reconstructions. Expanding the definition of preservation to include such reference data is therefore essential in environments where evidence may need to be interpreted long after the fact.

> **Research Gap 3**. *There is little work on the persistence of artifacts, and determining if the absence of data is due to configuration, tampering, or simply the passing of time. Work in this area could reduce this aspect of uncertainty within the model and process, and provide practical advice on the temporal boundaries of useful preservation periods.*

### 9.3. Q3: Timeline

This aspect of event reconstruction has received the most attention and many articles and concepts have been discussed.

**Continuous updates/improvement to timestamp extraction:** Files and formats containing timestamps are subject to change. Ongoing research that tracks these changes and uncovers new timestamp sources provides the foundational data necessary. This means ongoing 'artifact research' (as defined by Breitinger et al. (2024)) is critical.

**Integration of non-explicit timing information:** Dreier et al. (2024) discussed implicit timing (e.g., ordering of log file entries) to detect inconsistencies in an automated way. A second possibility is digital stratigraphy, as defined by Casey (2018), and further implemented in Schneider et al. (2024), which is a method that takes advantage of file systems and the behavior of their allocation algorithms. By analyzing the logical position of files on a disk, investigators can infer potential events, provided they understand how the file system allocates those files. This knowledge enables the reconstruction of hypothetical sequences of events based on file placement. These are still early implementations, and additional work is needed to evaluate more variations in environments, file systems, drivers, and behavior patterns.

**Timeline representation:** Timelines are mostly flat, i.e., textual files in chronological order. The community should explore alternatives. For instance, an ontology-based approach improves event reconstruction by providing a structured and formal representation of data, which helps standardize and automate the analysis process (Bhandari and Jusas, 2020). An ontology captures the semantic relationships between events, objects, and subjects, allowing investigators to infer new facts, identify correlations between events, and visualize data more effectively (Chabot et al., 2015b; Turnbull and Randhawa, 2015). We should also reconsider visualizing timelines, moving beyond the frequently used basic bar charts counting the number of events within defined timeframes, and exploring AR or VR.

**Automated timeline verification:** Willassen (2008c) introduced a hypothesis-based approach where investigators create clock hypotheses to model historical clock values and test their consistency with timestamp evidence. Vanini et al. (2024) suggested using time anchors (i.e., artifacts that include internal and external timestamps) and looking for anomalies. Research efforts need to continue to build verification methods that allow us to identify whether the timeline is out-of-sequence (irregularities found) or likely correct.

**Tamper detection:** Galhuber and Luh (2021) found that timestamp forgery tools may introduce detectable changes, such as reducing timestamp accuracy from nanoseconds to seconds. Among the tools they evaluated, only one was capable of modifying the full range of file system timestamps on Windows. Andrade (2020) noted that $FN timestamps are typically modified only by the Windows kernel and are generally unaffected by anti-forensic timestomping tools, offering an example of a timestamp that is harder to manipulate during event reconstruction. Jang et al. (2016) presented a method to detected time manipulation in NTFS file system. More general experiments as conducted by Schneider et al. (2020, 2022); Vanini et al. (2025) show that the probability of detecting it is high, especially when it concerns file metadata. One reason is that it is difficult to forge a timestamp without causing subsequent inconsistencies. While some progress has been made in detecting tampering, this area still requires further exploration and automation. Ideally, a tool should be capable of analyzing a timeline and automatically highlighting all potential tampering events.

> **Research Gap 4.** *Advances in timeline generation research are still needed in multiple areas: from artifact research, integration of non-timestamp-based timing information, visualization of timelines, and detecting inconsistencies and tampering.*

### 9.4. Q4: Analysis and investigative conclusions

This includes the timeline analysis which bridges Q3 and Q4 since it may revisited as part of Q4 hypothesis testing.

**Timeline analysis:** Efforts focus on methods to reduce and manage data, including techniques for filtering, labeling, and aggregating data. Flagging entries that match certain criteria can be performed, or more complex approaches such as discussed by Hargreaves and Patterson (2012); Studiawan et al. (2020b) where patterns of events are bundled to provide multiple entries that support an event reconstruction. This reduces large timelines to more manageable sets of interesting events, but as they are inherently a reduced set, switching back to the lower-level entry view is an important feature to retain to see inferred events in context and show provenance of the reconstructed event. A limitation discussed by Hargreaves and Patterson (2012) is the need to manually create the patterns that need to be matched based on research and experience. Better centralized documentation of the expected changes from sets of actions in different environments, similar to Casey et al. (2022); Grajeda et al. (2018) and integration into a standard timeline analysis tool would make timeline analysis more accessible.

Visualization is also a vital additional layer of abstraction to help make sense of the large amounts of data, and can be a valuable tool to assist with analysis, e.g., to support timeline-based cross drive analysis (Patterson and Hargreaves, 2012).

An increased availability of ground truth data sets with annotation of the actions carried out would assist with developing analysis plugins for tools (Grajeda et al., 2017). Automated event inference, either using machine learning, or through automation in digital forensic experimentation to carry out actions and record the resulting traces may help with this.

**Artifact reliability:** If the timeline contains conflicting information i.e., at least two artifacts provide conflicting information, a resolution is needed. Automation in identifying accurate artifacts would be advantageous. One possibility is to compare artifacts and assess their reliability, e.g., the ease of manipulating an artifact (Vanini et al., 2025).

Hargreaves and Patterson (2012) began work on handling conflicting artifacts, where each inferred high-level event was assigned a series of expected artifacts. On a match, the supporting *and contradictory* timeline entries were stored within the inferred event, highlighting entries that were expected but absent, forming the basis for the evaluation of reliability assessment. Casey (2011) discusses the number of independent sources and their resistance to tampering as part of the C-Scale, but if this were to be more strictly quantified, e.g., with Bayesian networks for example (Kwan et al., 2008), in terms of assigning weight to expected artifacts, other factors may have an impact. For example 'artifact longevity', i.e., how long an artifact is known to persist may allow appropriate weight to be given to the absence of specific, expected, hypothesis-supporting information. It remains unclear how appropriate precise numerical assessments in event reconstruction are.

**AI integration:** The use of AI for digital forensics is becoming more common (Du et al., 2020a; Jarrett and Choo, 2021). AI can help analyze and identify digital evidence (Henseler and van Beek, 2023; Sreya and Wadhwa, 2023) or aid investigators in writing forensic reports (Michelet and Breitinger, 2024). As discussed by Scanlon et al. (2023), LLMs may help with event analysis, such as suspicious activities or attack identification. However, they may hallucinate when responding to investigator questions. Others have tried to apply AI techniques to accelerate the process, e.g., by searching for anomalies (Studiawan et al., 2017; Studiawan and Sohel, 2021) or relevant artifacts (Du et al., 2020b; Marková et al., 2022). Future work should focus on evaluating and validating this new technology for forensic purposes.

**Natural Language Processing (NLP) integration:** NLP may support timeline analysis as each event is represented by a descriptive message. These messages contain valuable information that can be extracted and analyzed. By applying traditional NLP techniques, such as sentiment analysis (Silalahi et al., 2023c; Studiawan et al., 2020b), named entity recognition (Silalahi et al., 2023a, 2023b; Studiawan et al., 2023), and information extraction, researchers can derive insights. For future research, there is potential to explore other NLP methods to enhance the field. For instance, topic modeling and dependency parsing could be employed to gain deeper insights into events and establish relationships between them.

**Process mining:** Event reconstruction is a common task in process mining (Weijters and van der Aalst, 2001; Jürgensen, 2021), though it is typically applied to business process logs (Nguyen and Comuzzi, 2019). However, the domain faces similar challenges. For example, Dixit et al. (2018) describe a set of timestamp-based indicators for identifying event ordering imperfections in logs and present a method for resolving these issues using domain knowledge. Therefore, future research could explore various process mining techniques (van der Aalst, 2016) for forensic event reconstruction.

**Training and education:** Specialized training and continuous education play a key role in ensuring investigators can handle complex cases and maintain the admissibility of evidence in court (Jahankhani & Hosseinian-far, 2014). However, cognitive biases and human errors can impact the integrity of findings, but some techniques can be used to mitigate this, e.g., collaborative approaches, such as the 4-eye principle, where at least two individuals review the findings. More research is needed to explore how collaborative techniques and advanced decision-support systems, including AI-assisted tools, can further minimize human errors and biases, ensuring more reliable and transparent event reconstruction processes.

> **Research Gap 5.** *The challenge of performing efficient and effective timeline analysis remains. Handling the volume of extracted timestamps in an effective way is needed (Q3/4), which could include technological solutions such as performance improvements or AI based filtering, but also process changes, where the 'extract everything' model needs research to ensure it is still the most appropriate approach.*

> **Research Gap 6.** *Automation is likely the only practical way to handle the challenge of inferring events at scale (Q4), but how to handle the*

*practical research challenge of automated inference of events from timeline entries that are subject to operating system, application, and environmental changes earlier on in the process (Q1,Q2) is challenging.*

**Research Gap 7.** *Ensuring and communicating a clear delineation between extracted timestamp values as facts, and inferred events as working hypothesis, in both research and in forensic tooling (Q4), requires work from digital forensic scientists, and potentially UX experts to clearly communicate residual uncertainty.*

## 10. Conclusions

Event reconstruction is a critical part of the digital forensic process, yet the process and terminology are vague and inconsistent. This work has shown that this mixture of terms can be unified and as a result, a systematic organization of issues associated with timeline-based event reconstruction w compiled. When an event reconstruction is completed, these potential issues can be considered and evaluated as to whether they may have influenced the result of the reconstruction. Aside from practical uses, it has also allowed clear future directions in event reconstruction research to be identified.

While some of these identified challenges will be obvious to seasoned investigators, there is a need within digital forensics to formalize definitions and make currently tacit knowledge explicit. This provides the foundation for more formal and potentially future quantitative evaluation of the trustworthiness or indeed reliability of reconstructed events in a digital forensic investigation.

### Disclosure of AI-assisted writing tools

Some authors utilized ChatGPT-4 to assist in revising, condensing text, and correcting grammatical errors, typos, and awkward phrasing. All AI-generated suggestions were carefully reviewed and modified as necessary to ensure they aligned with the authors' intended meaning before being incorporated into this paper.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

We acknowledge Eoghan Casey for the comments and feedback. The authors also thank Céline Vanini for the initial diagram and discussions.

## References

van der Aalst, W., 2016. Data science in action. In: Process Mining: Data Science in Action. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–23. https://doi.org/10.1007/978-3-662-49851-4_1.

Adderley, N., Peterson, G., 2020. Interactive temporal digital forensic event analysis. In: Peterson, G., Shenoi, S. (Eds.), *Advances In Digital Forensics XVI* IFIP Advances in Information and Communication Technology. Springer International Publishing, Cham, pp. 39–55. https://doi.org/10.1007/978-3-030-56223-6_3.

Adedayo, O.M., Olivier, M.S., 2015. Ideal log setting for database forensics reconstruction. Digit. Invest. 12, 27–40.

Alqahtany, S., Clarke, N., Furnell, S., Reich, C., 2016. A forensic acquisition and analysis system for IaaS. Clust. Comput. 19, 439–453. https://doi.org/10.1007/s10586-015-0509-x.

Amato, F., Cozzolino, G., Mazzeo, A., Mazzocca, N., 2017. Correlation of digital evidences in forensic investigation through semantic technologies. In: 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 668–673. https://doi.org/10.1109/WAINA.2017.4.

Andrade, R., 2020. Expose Evidence of Timestomping with the Ntfs Timestamp Mismatch Artifact. URL: https://www.magnetforensics.com/blog/expose-evidence-of-timest omping-with-the-ntfs-timestamp-mismatch-artifact-in-magnet-axiom-4-4/.

Arshad, H., Jantan, A.B., Abiodun, O.I., 2018. Digital forensics: review of issues in scientific validation of digital evidence. Journal of Information Processing Systems 14, 346–376.

Bang, J., Yoo, B., Kim, J., Lee, S., 2009. Analysis of time information for digital investigation. In: 2009 Fifth International Joint Conference on INC, IMS and IDC. IEEE, pp. 1858–1864.

Batten, L., Pan, L., Khan, N., 2012. Hypothesis generation and testing in event profiling for digital forensic investigations. Int. J. Digital Crime Forensics (IJDCF) 4, 1–14. https://doi.org/10.4018/jdcf.2012100101.

Battistoni, R., Di Pietro, R., Lombardi, F., 2016. Cure—towards enforcing a reliable timeline for cloud forensics: model, architecture, and experiments. Comput. Commun. 91, 29–43.

Becker, D., Rabenseifner, R., Wolf, F., 2008. Implications of non-constant clock drifts for the timestamps of concurrent events. In: 2008 IEEE International Conference on Cluster Computing, pp. 59–68.

Berggren, J., Gudjonsson, K., Jäger, A., et al., 2024. Timesketch: collaborative forensic timeline analysis. https://github.com/google/timesketch.

Bhandari, S., Jusas, V., 2020. An ontology based on the timeline of Log2timeline and Psort using abstraction approach in digital forensics. Symmetry 12, 642. https://doi.org/10.3390/sym12040642. URL: https://www.mdpi.com/2073-8994/12/4/642. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

Bhat, W.A., AlZahrani, A., Wani, M.A., 2021. Can computer forensic tools be trusted in digital investigations? Sci. Justice 61, 198–203.

Boyd, C., Forster, P., 2004. Time and date issues in forensic computing—a case study. Digit. Invest. 1, 18–23.

Breitinger, F., Hilgert, J.-N., Hargreaves, C., Sheppard, J., Overdorf, R., Scanlon, M., 2024. Dfrws eu 10-year review and future directions in digital forensic research. Forensic Sci. Int.: Digit. Invest. 48, 301685.

Buchholz, F., Tjaden, B., 2007. A brief study of time. Digit. Invest. 4, 31–42. https://doi.org/10.1016/j.diin.2007.06.004.

Buchholz, F.P., Falk, C., 2005. Design and implementation of zeitline: a forensic timeline editor. In: DFRWS.

Carrier, B., Spafford, E., 2004a. Defining event reconstruction of a digital crime scene. J. Forensic Sci. 49, 1291–1298. https://doi.org/10.1520/JFS2004127.

Carrier, B., Spafford, E., 2004b. An event-based digital forensic investigation framework. In: Proceedings of the the Digital Forensic Research Conference, pp. 1–12.

Carrier, B.D., 2006. A Hypothesis-Based Approach to Digital Forensic Investigations. Ph. D. thesis Purdue University.

Casey, E., 2011. Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet, third ed. Academic Press, Waltham, MA.

Casey, E., 2018. Digital stratigraphy: contextual analysis of file system traces in forensic science. J. Forensic Sci. 63, 1383–1391. https://doi.org/10.1111/1556-4029.13722. Number: 5.

Casey, E., 2020. Standardization of forming and expressing preliminary evaluative opinions on digital evidence. Forensic Sci. Int.: Digit. Invest. 32, 200888. https://doi.org/10.1016/j.fsidi.2019.200888.

Casey, E., Nguyen, L., Mates, J., Lalliss, S., 2022. Crowdsourcing forensics: creating a curated catalog of digital forensic artifacts. J. Forensic Sci. 67, 1846–1857. https://doi.org/10.1111/1556-4029.15053._eprint. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1556-4029.15053.

Chabot, Y., Bertaux, A., Nicolle, C., Kechadi, M.-T., 2014. A complete formalized knowledge representation model for advanced digital forensics timeline analysis. Digit. Invest. 11, S95–S105. https://doi.org/10.1016/j.diin.2014.05.009.

Chabot, Y., Bertaux, A., Nicolle, C., Kechadi, M.-T., 2015a. Event reconstruction: a state of the art. In: Cruz-Cunha, M.M., Portela, I.M., Piekarz, A. (Eds.), *Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance:* Advances in Digital Crime, Forensics, and Cyber Terrorism. IGI Global, p. 15. https://doi.org/10.4018/978-1-4666-6324-4.

Chabot, Y., Bertaux, A., Nicolle, C., Kechadi, T., 2015b. An ontology-based approach for the reconstruction and analysis of digital incidents timelines. Digit. Invest. 15, 83–100.

Choi, H., Lee, S., Jeong, D., 2021. Forensic recovery of SQL server database: practical approach. IEEE Access 9, 14564–14575.

Conlan, K., Baggili, I., Breitinger, F., 2016. Anti-forensics: furthering digital forensic science through a new extended, granular taxonomy. Digit. Invest. 18, S66–S75. https://doi.org/10.1016/j.diin.2016.04.006.

Debinski, M., Breitinger, F., Mohan, P., 2019. Timeline2GUI: a Log2Timeline CSV parser and training scenarios. Digit. Invest. 28, 34–43. https://doi.org/10.1016/j.diin.2018.12.004.

Dixit, P.M., Suriadi, S., Andrews, R., Wynn, M.T., ter Hofstede, A.H., Buijs, J.C., van der Aalst, W.M., 2018. Detection and interactive repair of event ordering imperfection in process logs. In: Advanced Information Systems Engineering: 30th International Conference, CAiSE 2018, Tallinn, Estonia, June 11-15, 2018, Proceedings, vol. 30. Springer, pp. 274–290.

Dreier, L.M., Vanini, C., Hargreaves, C.J., Breitinger, F., Freiling, F., 2024. Beyond timestamps: integrating implicit timing information into digital forensic timelines. Forensic Sci. Int.: Digit. Invest. 49, 301755. https://doi.org/10.1016/j.fsidi.2024.301755.

Du, X., Hargreaves, C., Sheppard, J., Anda, F., Sayakkara, A., Le-Khac, N.-A., Scanlon, M., 2020a. SoK: exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation. In: Proceedings of the 15th International Conference on Availability, Reliability and Security, pp. 1–10.

Du, X., Le, Q., Scanlon, M., 2020b. Automated artefact relevancy determination from artefact metadata and associated timeline events. In: 2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security). IEEE, pp. 1–8.

Fernández-Fuentes, X., Pena, T.F., Cabaleiro, J.C., 2022. Digital forensic analysis methodology for private browsing: firefox and chrome on linux as a case study. Comput. Secur. 115, 102626.

Galhuber, M., Luh, R., 2021. Time for truth: forensic analysis of NTFS timestamps. In: *Proceedings of the 16th International Conference on Availability, Reliability and Security* ARES 21. Association for Computing Machinery, New York, NY, USA, pp. 1–10. https://doi.org/10.1145/3465481.3470016.

Gladyshev, P., Patel, A., 2004. Finite state machine approach to digital event reconstruction. Digit. Invest. 1, 130–149. https://doi.org/10.1016/j.diin.2004.03.001.

Gómez, R., Herrerias, J., Mata, E., 2005. Using lamport's logical clocks to consolidate log files from different sources. In: International Workshop on Innovative Internet Community Systems. Springer, pp. 126–133.

Grajeda, C., Breitinger, F., Baggili, I., 2017. Availability of datasets for digital forensics–and what is missing. Digit. Invest. 22, S94–S105.

Grajeda, C., Sanchez, L., Baggili, I., Clark, D., Breitinger, F., 2018. Experience constructing the artifact genome project (agp): managing the domain's knowledge one artifact at a time. Digit. Invest. 26, S47–S58.

Gruber, J., Hargreaves, C.J., Freiling, F.C., 2023. Contamination of digital evidence: understanding an underexposed risk. Forensic Sci. Int.: Digit. Invest. 44, 301501. https://doi.org/10.1016/j.fsidi.2023.301501.

Gudhjónsson, K., 2010. Mastering the Super Timeline with Log2timeline. SANS Institute.

Hampton, N., Baig, Z.A., 2016. Timestamp analysis for quality validation of network forensic data. In: Network and System Security: 10th International Conference, NSS 2016, Taipei, Taiwan, September 28-30, 2016, Proceedings 10. Springer, pp. 235–248.

Han, J., Kim, J., Lee, S., 2020. 5w1h-based expression for the effective sharing of information in digital forensic investigations. arXiv, 15711 preprint arXiv:2010.

Hargreaves, C., van Beek, H., Casey, E., 2025. Solve-it: a proposed digital forensic knowledge base inspired by mitre att&ck. Forensic Sci. Int.: Digit. Invest. 52, 301864.

Hargreaves, C., Breitinger, F., Dowthwaite, L., Webb, H., Scanlon, M., 2024a. Dfpulse: the 2024 digital forensic practitioner survey. Forensic Sci. Int.: Digit. Invest. 51, 301844.

Hargreaves, C., Nelson, A., Casey, E., 2024b. An abstract model for digital forensic analysis tools-a foundation for systematic error mitigation analysis. Forensic Sci. Int.: Digit. Invest. 48, 301679. https://doi.org/10.1016/j.fsidi.2023.301679.

Hargreaves, C., Patterson, J., 2012. An automated timeline reconstruction approach for digital forensic investigations. Digit. Invest. 9, S69–S79. https://doi.org/10.1016/j.diin.2012.05.006.

Hargreaves, C.J., 2009. Assessing the Reliability of Digital Evidence from Live Investigations Involving Encryption. Ph.D. thesis Cranfield University, UK.

Harichandran, V.S., Walnycky, D., Baggili, I., Breitinger, F., 2016. Cufa: a more formal definition for digital forensic artifacts. Digit. Invest. 18, S125–S137.

Henderson, G., 2009. *A Categorization of Computer Clocks*. Technical Report Department of Computer Science. James Madison University.

Henseler, H., van Beek, H., 2023. Chatgpt as a copilot for investigating digital evidence. In: Proceedings of the Third International Workshop on Artificial Intelligence and Intelligent Assistance for Legal Professionals in the Digital Workplace, pp. 58–69.

Herman, M., Herman, M., Iorga, M., Salim, A.M., Jackson, R.H., Hurst, M.R., Leo, R., Lee, R., Landreville, N.M., Mishra, A.K., Wang, Y., 2020. *Nist cloud computing forensic science challenges*. US Department of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, USA.

Horsman, G., 2018. "i couldn't find it your honour, it mustn't be there!"–tool errors, tool limitations and user error in digital forensics. Sci. Justice 58, 433–440.

Horsman, G., 2019. Raiders of the lost artefacts: championing the need for digital forensics research. Forensic Sci. Int.: Report 1, 100003.

Inglot, B., Liu, L., 2014. Enhanced timeline analysis for digital forensic investigations. Inf. Secur. J. A Glob. Perspect. 23, 32–44.

Jahankhani, H., Hosseinian-far, A., 2014. Digital forensics education, training and awareness. In: Cyber Crime and Cyber Terrorism Investigator's Handbook. Elsevier, pp. 91–100. https://doi.org/10.1016/B978-0-12-800743-3.00008-6.

Jang, D.-i., Ahn, G.-J., Hwang, H., Kim, K., 2016. Understanding anti-forensic techniques with timestamp manipulation. In: 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI). IEEE, pp. 609–614.

Jaquet-Chiffelle, D.-O., Casey, E., 2021. A formalized model of the Trace. Forensic Sci. Int. 327, 110941. https://doi.org/10.1016/j.forsciint.2021.110941.

Jarrett, A., Choo, K.-K.R., 2021. The impact of automation and artificial intelligence on digital forensics. Wiley Interdisciplinary Reviews: Forensic Sci. 3, e1418.

Jinad, R., Gupta, K., Simsek, E., Zhou, B., 2024. Bias and fairness in software and automation tools in digital forensics. J. Surveill. Secur. Saf 5, 19–35.

Joseph, A., Singh, K.J., 2019. Digital forensics in distributed environment. In: Cloud Security: Concepts, Methodologies, Tools, and Applications. IGI Global, pp. 1157–1177.

Jürgensen, J.P., 2021. Trace reconstruction in system logs for processing with process mining. Procedia Comput. Sci. 180, 352–357.

Kaart, M., Laraghy, S., 2014. Android forensics: interpretation of timestamps. Digit. Invest. 11, 234–248. https://doi.org/10.1016/j.diin.2014.05.001.

Kang, J., Lee, S., Lee, H., 2013. A digital forensic framework for automated user activity reconstruction. In: Information Security Practice and Experience: 9th International Conference, ISPEC 2013, Lanzhou, China, May 12-14, 2013. Proceedings, vol. 9. Springer, pp. 263–277.

Karie, N.M., Venter, H.S., 2015. Taxonomy of challenges for digital forensics. J. Forensic Sci. 60, 885–893.

Kassin, S.M., Dror, I.E., Kukucka, J., 2013. The forensic confirmation bias: problems, perspectives, and proposed solutions. Journal of applied research in memory and cognition 2, 42–52.

Kebande, V.R., Venter, H.S., 2018. Novel digital forensic readiness technique in the cloud environment. Aust. J. Forensic Sci. 50, 552–591. https://doi.org/10.1080/00450618.2016.1267797.

Khan, M., Chatwin, C.R., Young, R.C., 2007. A framework for post-event timeline reconstruction using neural networks. Digit. Invest. 4, 146–157.

Khan, M., Wakeman, I., 2006. Machine learning for post-event timeline reconstruction. In: First Conference on Advances in Computer Security and Forensics. Citeseer, Liverpool, UK.

Khan, S., Gani, A., Wahab, A.W.A., Bagiwa, M.A., Shiraz, M., Khan, S.U., Buyya, R., Zomaya, A.Y., 2016. Cloud log forensics: foundations, state of the art, and future directions. ACM Comput. Surv. 49, 1–42. https://doi.org/10.1145/2906149.

Kiernan, J., Terzi, E., 2009. Eventsummarizer: a tool for summarizing large event sequences. In: Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology, pp. 1136–1139.

Kłos, M., El Fray, I., 2020. Securing event logs with blockchain for iot. In: International Conference on Computer Information Systems and Industrial Management. Springer, pp. 77–87. https://doi.org/10.1007/978-3-030-47679-3_7.

Kwan, M., Chow, K.-P., Law, F., Lai, P., 2008. Reasoning about evidence using bayesian networks. In: Advances in Digital Forensics IV, vol. 4. Springer, pp. 275–289.

Kälber, S., Dewald, A., Freiling, F.C., 2013. Forensic application-fingerprinting based on file system metadata. In: 2013 Seventh International Conference on IT Security Incident Management and IT Forensics, pp. 98–112. https://doi.org/10.1109/IMF.2013.20.

Latzo, T., Freiling, F., 2019. Characterizing the limitations of forensic event reconstruction based on log files. In: 2019 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), pp. 466–475. https://doi.org/10.1109/TrustCom/BigDataSE.2019.00069 iSSN: 2324-9013.

Lee, H.C., Palmbach, T., Miller, M.T., 2001. Henry Lee's Crime Scene Handbook. Academic Press.

Levett, C.P., Jhumka, A., Anand, S.S., 2010. Towards event ordering in digital forensics. In: Proceedings of the 12th ACM Workshop on Multimedia and Security, pp. 35–42.

Losavio, M., Pastukov, P., Polyakova, S., 2015. Cyber black box/event data recorder: legal and ethical perspectives and challenges with digital forensics. Journal of Digital Forensics, Security and Law 10, 4.

Lyle, J.R., Guttman, B., Butler, J.M., Sauerwein, K., Reed, C., Lloyd, C.E., 2022. Digital Investigation Techniques: A NIST Scientific Foundation Review. Technical Report National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8354-draft.

Malhotra, A., Cohen, I.E., Brakke, E., Goldberg, S., 2015. Attacking the Network Time Protocol. Cryptology ePrint Archive.

Manral, B., Somani, G., Choo, K.-K.R., Conti, M., Gaur, M.S., 2019. A systematic survey on cloud forensics challenges, solutions, and future directions. ACM Comput. Surv. 52, 1–38.

Marangos, N., Rizomiliotis, P., Mitrou, L., 2016. Time synchronization: pivotal element in cloud forensics. Secur. Commun. Network. 9, 571–582.

Marková, E., Sokol, P., Kováčová, K., 2022. Detection of relevant digital evidence in the forensic timelines. In: 2022 14th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). IEEE, pp. 1–7.

Marrington, A., Baggili, I., Mohay, G., Clark, A., 2011. CAT Detect (Computer Activity Timeline Detection): a tool for detecting inconsistency in computer activity timelines. Digit. Invest. 8, S52–S61.

Marrington, A., Mohay, G., Clark, A., Morarji, H., 2007. Event-based Computer Profiling for the Forensic Reconstruction of Computer Activity. AusCERT 2007, IT-Security: Finding the Balance, pp. 71–87.

Metz, J., Gudjonsson, K., White, D., et al., 2024. log2timeline Plaso: super timeline all the things. https://github.com/log2timeline/plaso.

Michelet, G., Breitinger, F., 2024. Chatgpt, llama, can you write my report? an experiment on assisted digital forensics reports written using (local) large language models. Forensic Sci. Int.: Digit. Invest. 48, 301683.

MITRE, 2023. Impair Defenses. https://attack.mitre.org/techniques/T1562/.

Mohammed, H., Clarke, N., Li, F., 2016. An automated approach for digital forensic analysis of heterogeneous big data. Journal of Digital Forensics, Security and Law 11, 9.

Neale, C., 2023. Fool me once: a systematic review of techniques to authenticate digital artefacts. Forensic Sci. Int.: Digit. Invest. 45, 301516. https://doi.org/10.1016/j.fsidi.2023.301516.

Neale, C., Kennedy, I., Price, B., Yu, Y., Nuseibeh, B., 2022. The case for zero trust digital forensics. Forensic Sci. Int.: Digit. Invest. 40, 301352.

Neasbitt, C., Perdisci, R., Li, K., Nelms, T., 2014. Clickminer: towards forensic reconstruction of user-browser interactions from network traces. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1244–1255.

Nguyen, H.T.C., Comuzzi, M., 2019. Event log reconstruction using autoencoders. In: Service-Oriented Computing–ICSOC 2018 Workshops. Springer, pp. 335–350.

Nordvik, R., Axelsson, S., 2022. It is about time–do exfat implementations handle timestamps correctly? Forensic Sci. Int.: Digit. Invest. 42, 301476.

Oh, J., Lee, S., Hwang, H., 2022. Forensic recovery of file system metadata for digital forensic investigation. IEEE Access 10, 111591–111606.

Osborne, G., Turnbull, B., 2009. Enhancing computer forensics investigation through visualisation and data exploitation. In: 2009 International Conference on Availability, Reliability and Security. IEEE, pp. 1012–1017.

Palmbach, D., Breitinger, F., 2020. Artifacts for detecting timestamp manipulation in NTFS on Windows and their reliability. Forensic Sci. Int.: Digit. Invest. 32, 300920. https://doi.org/10.1016/j.fsidi.2020.300920.

Patterson, J., Hargreaves, C.J., 2012. The Potential for cross-drive analysis using automated digital forensic timelines. https://dspace.lib.cranfield.ac.uk/handle/1826/8088.Accepted:2014-01-23T05:01:12Z.

Quick, D., Choo, K.-K.R., 2014. Impacts of increasing volume of digital forensic data: a survey and future research challenges. Digit. Invest. 11, 273–294.

Raghavan, S., Saran, H., 2013. Unitime: timestamp interpretation engine for developing unified timelines. In: 2013 8th International Workshop on Systematic Approaches to Digital Forensics Engineering (SADFE). IEEE, pp. 1–7.

Raju, B.K, Gosala, N.B., Geethakumari, G., 2017. Closer: applying aggregation for effective event reconstruction of cloud service logs. In: Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication, pp. 1–8.

Reddy, K., Venter, H.S., 2013. The architecture of a digital forensic readiness management system. Comput. Secur. 32, 73–89. https://doi.org/10.1016/j.cose.2012.09.008.

Ribaux, O., 2014. Police scientifique: le renseignement par la trace. Sciences forensiques. Presses polytechniques et universitaires romandes, Lausanne.

Ribaux, O., 2023. De la police scientifique à la traçologie: le renseignement par la trace. EPFL Press.

Rivera-Ortiz, F., Pasquale, L., 2019. Towards automated logging for forensic-ready software systems. In: 2019 IEEE 27th International Requirements Engineering Conference Workshops (REW). IEEE, pp. 157–163. https://doi.org/10.1109/REW.2019.00033.

Roux, C., Bucht, R., Crispino, F., De Forest, P., Lennard, C., Margot, P., Miranda, M.D., NicDaeid, N., Ribaux, O., Ross, A., Willis, S., 2022. The Sydney declaration – Revisiting the essence of forensic science through its fundamental principles. Forensic Sci. Int. 332, 111182. https://doi.org/10.1016/j.forsciint.2022.111182.

Sachowski, J., 2019. Implementing Digital Forensic Readiness: from Reactive to Proactive Process. CRC Press. https://doi.org/10.1016/C2015-0-00701-8.

Sandvik, J.-P., Franke, K., Årnes, A., 2021. Towards a generic approach of quantifying evidence volatility in resource constrained devices. Digital Forensic Investigation of Internet of Things (IoT) Devices 21–45.

Sandvik, J.-P., Årnes, A., 2018. The reliability of clocks as digital evidence under low voltage conditions. Digit. Invest. 24, S10–S17. https://doi.org/10.1016/j.diin.2018.01.003.

Scanlon, M., Breitinger, F., Hargreaves, C., Hilgert, J.-N., Sheppard, J., 2023. ChatGPT for digital forensic investigation: the good, the bad, and the unknown. Forensic Sci. Int.: Digit. Invest. 46, 301609.

Schatz, B., Mohay, G., Clark, A., 2006. A correlation method for establishing provenance of timestamps in digital evidence. Digit. Invest. 3, 98–107. https://doi.org/10.1016/j.diin.2006.06.009.

Schneider, J., Düsel, L., Lorch, B., Drafz, J., Freiling, F., 2022. Prudent design principles for digital tampering experiments. Forensic Sci. Int.: Digit. Invest. 40, 301334. https://doi.org/10.1016/j.fsidi.2022.301334.

Schneider, J., Eichhorn, M., Dreier, L.M., Hargreaves, C., 2024. Applying digital stratigraphy to the problem of recycled storage media. Forensic Sci. Int.: Digit. Invest. 49, 301761.

Schneider, J., Wolf, J., Freiling, F., 2020. Tampering with digital evidence is hard: the case of main memory images. Forensic Sci. Int.: Digit. Invest. 32, 300924. https://doi.org/10.1016/j.fsidi.2020.300924.

Schuster, A., 2007. Introducing the microsoft vista event log file format. Digit. Invest. 4, 65–72.

Silalahi, S., Ahmad, T., Studiawan, H., 2023a. Dfler: drone flight log entity recognizer to support forensic investigation on drone device. Software Impacts 15, 100457. https://doi.org/10.1016/j.simpa.2022.100457.

Silalahi, S., Ahmad, T., Studiawan, H., 2023b. Transformer-based named entity recognition on drone flight logs to support forensic investigation. IEEE Access 11, 3257–3274. https://doi.org/10.1109/ACCESS.2023.3234605.

Silalahi, S., Ahmad, T., Studiawan, H., 2023c. Transformer-based sentiment analysis for anomaly detection on drone forensic timeline. In: 2023 11th International Symposium on Digital Forensics and Security (ISDFS). IEEE, pp. 1–6. ISDFS58141.2023.10131749.

Soltani, S., Hosseini Seno, S.A., sadoghi yazdi, H., 2019. Event reconstruction using temporal pattern of file system modification. IET Inf. Secur. 13. https://doi.org/10.1049/iet-ifs.2018.5209.

Soltani, S., Seno, S.A.H., 2017. A survey on digital evidence collection and analysis. In: 2017 7th International Conference on Computer and Knowledge Engineering (ICCKE). IEEE, pp. 247–253.

Soltani, S., Seno, S.A.H., 2019. A formal model for event reconstruction in digital forensic investigation. Digit. Invest. 30, 148–160. https://doi.org/10.1016/j.diin.2019.07.006.

Song, S., Cao, Y., Wang, J., 2016. Cleaning timestamps with temporal constraints. Proceedings of the VLDB Endowment 9, 708–719.

Spichiger, H., Adelstein, F., 2025. Preserving meaning of evidence from evolving systems. Forensic Sci. Int.: Digit. Invest. 52, 301867. https://doi.org/10.1016/j.fsidi.2025.301867. URL: https://www.sciencedirect.com/science/article/pii/S266628172500006X. DFRWS EU 2025 - Selected Papers from the 12th Annual Digital Forensics Research Conference Europe.

Sreya, E., Wadhwa, M., 2023. Enhancing digital investigation: leveraging chatgpt for evidence identification and analysis in digital forensics. In: 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). IEEE, pp. 733–738.

Stevens, M.W., 2004. Unification of relative time frames for digital forensics. Digit. Invest. 1, 225–239. https://doi.org/10.1016/j.diin.2004.07.003.

Studiawan, H., 2023. Event abstration in a forensic timeline. In: International Conference for Information and Communication Technologies. Springer, pp. 119–129.

Studiawan, H., Hasan, M.F., Pratomo, B.A., 2023. Rule-based entity recognition for forensic timeline. In: 2023 Conference on Information Communications Technology and Society (ICTAS). IEEE, pp. 1–6.

Studiawan, H., Payne, C., Sohel, F., 2017. Graph clustering and anomaly detection of access control log for forensic purposes. Digit. Invest. 21, 76–87.

Studiawan, H., Sohel, F., 2021. Anomaly detection in a forensic timeline with deep autoencoders. J. Inf. Secur. Appl. 63, 103002.

Studiawan, H., Sohel, F., Payne, C., 2019. A survey on forensic investigation of operating system logs. Digit. Invest. 29, 1–20. https://doi.org/10.1016/j.diin.2019.02.005.

Studiawan, H., Sohel, F., Payne, C., 2020a. Automatic event log abstraction to support forensic investigation. In: Proceedings of the Australasian Computer Science Week Multiconference, pp. 1–9.

Studiawan, H., Sohel, F., Payne, C., 2020b. Sentiment analysis in a forensic timeline with deep learning. IEEE Access 8, 60664–60675. https://doi.org/10.1109/ACCESS.2020.2983435.

Thierry, A., Müller, T., 2022. A systematic approach to understanding MACB timestamps on Unix-like systems. Forensic Sci. Int.: Digit. Invest. 40, 301338.

Turnbull, B., Randhawa, S., 2015. Automated event and social network extraction from digital evidence sources with ontological mapping. Digit. Invest. 13, 94–106. https://doi.org/10.1016/j.diin.2015.04.004.

Vanini, C., Breitinger, F., Hargreaves, C., 2023. A discussion of sources and quality/reliability of events for timelines. In: Presentation at the Digital Forensics Research Conference 2023 (Bonn, Germany).

Vanini, C., Gruber, J., Hargreaves, C., Benenson, Z., Freiling, F., Breitinger, F., 2025, April. Understanding Strategies and Challenges of Timestamp Tampering for Improved Digital Forensic Event Reconstruction. In: Proceedings of the Digital Forensics Doctoral Symposium, pp. 1–8.

Vanini, C., Hargreaves, C.J., van Beek, H., Breitinger, F., 2024. Was the clock correct? Exploring timestamp interpretation through time anchors for digital forensic event reconstruction. Forensic Sci. Int.: Digit. Invest. 49, 301759. https://doi.org/10.1016/j.fsidi.2024.301759.

VMware, 2008. Timekeeping in VMware Virtual Machines. https://www.cse.psu.edu/buu1/teaching/spring06/papers/vmware-timing.pdf.

Weijters, A., van der Aalst, W.M., 2001. Process mining: discovering workflow models from event-based data. In: Belgium-Netherlands Conf. On Artificial Intelligence.

Willassen, S., 2008a. Hypothesis-based investigation of digital timestamps. In: Ray, I., Shenoi, S. (Eds.), Advances in Digital Forensics IV IFIP — the International Federation for Information Processing, vol. 285. Springer US, Boston, MA, pp. 75–86. https://doi.org/10.1007/978-0-387-84927-0_7.

Willassen, S.Y., 2008b. Finding evidence of antedating in digital investigations. In: 2008 Third International Conference on Availability, Reliability and Security, pp. 26–32. https://doi.org/10.1109/ARES.2008.149.

Willassen, S.Y., 2008c. Timestamp evidence correlation by model based clock hypothesis testing. In: Proceedings of the 1st International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia and Workshop E-Forensics '08, pp. 1–6. Brussels, BEL: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Xu, W., Xu, D., 2022. Visualizing and reasoning about presentable digital forensic evidence with knowledge graphs. In: 2022 19th Annual International Conference on Privacy, Security & Trust (PST). IEEE, pp. 1–10.