



# **The Cyber-traceological Model: A Model-based View of the Cybercriminalistic Task**

By:  
Jan Gruber, Felix Freiling

From the proceedings of  
The Digital Forensic Research Conference  
**DFRWS APAC 2024**  
Oct 22-24, 2024

**DFRWS** is dedicated to the sharing of knowledge and ideas about digital forensics research. Ever since it organized the first open workshop devoted to digital forensics in 2001, DFRWS continues to bring academics and practitioners together in an informal environment.

As a non-profit, volunteer organization, DFRWS sponsors technical working groups, annual conferences and challenges to help drive the direction of research and development.

**<https://dfrws.org>**

# The Cyber-traceological Model: A Model-based View of the Cybercriminalistic Task

Jan Gruber, Felix Freiling\*

Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Martensstr. 3, 91058, Erlangen, Germany

---

## ARTICLE INFO

*Keywords:*  
cybercriminalistics  
digital investigations  
formal methods  
model-based approach

---

## ABSTRACT

The quantity and complexity of digital evidence pose significant challenges for solving both computer-enabled and core cybercrime cases. The very nature of digital systems renders traditional procedures of evidence collection and examination, such as a meticulous and systematic analysis of each and every potential trace, infeasible. To address this issue, we apply a model-based view to the cybercriminalistic task, which is comprised of the search for case-relevant hypotheses and the consequent identification of relevant traces to assess those previously identified hypotheses. To this end, we propose the *Cyber-traceological Model* helping to translate investigative questions to “relevant digital evidence” with which investigative hypotheses can be assessed. In the best case, we can use the Cyber-traceological Model to directly “compute” relevant digital evidence if a complete formal model of the system under investigation is already available. But even if such a model is not at hand, the Cyber-traceological Model can guide the search for relevant evidence in submodels, as we show in an example case of distributing prohibited multimedia data. Furthermore, we discuss the potential of model-based approaches in the field of forensic science in general, point out important research directions for shaping the emerging research discipline of cybercriminalistics, and ground it in formal methods.

---

## 1. Introduction

For the past 20 years, forensic scientists and criminalists alike have been confronted with a steep rise in the amount and complexity of digital data in criminal cases. The task in (digital) investigations is to pose apt investigative hypotheses and look at the expressive yet relevant traces. However, investigators face rapid technological advancements but have an incomplete understanding of cybercriminalistics (Gruber et al., 2022). This deficiency leads to a notable inefficiency in fighting cybercrime (Anderson et al., 2013) and often even to an inability to conduct resilient attribution (Berghele, 2017); hence, the present article tackles the fundamental but complicated question of how to conceptually find “relevant digital evidence”, i.e., enough telling traces to assess the hypotheses in question, and solve the criminalistic task eventually.

### 1.1. Motivation

It is extremely difficult to determine beforehand which traces will be decisive in solving a criminal case. Thus, investigators working on serious crime cases often employ a meticulous approach, striving for completeness when processing physical evidence. In doing so, they follow traces whose importance seems far-fetched at the time of processing but which may turn out later to be decisive. This “trawling net method” of investigation is often necessary because there is—apart from some experiential knowledge—no clear understanding of what traces are of critical relevance in the concrete case. Obviously, such an approach is rather resource-heavy.

The problems of the trawling net method are aggravated when digital evidence comes into play. This is due to the seemingly infinite amount of digital traces on even a simple system, the quickly increasing number of computing devices, and the ever-rising storage capacities of such devices. With digital evidence, it appears much easier to lose sight or at least lose focus on what can be relevant to a case. So given the wealth of available data and its increasing complexity (Carrier, 2003), the question of which evidence is relevant to prove which offense is as relevant as ever.

The advent of the vast amount of digital traces in criminal proceedings has, however, also highlighted the differences between physical and digital evidence: While the former is based on universally valid laws of nature—even when dealing with human-made analogous items—the latter is a result of human minds designing pieces of software instructing machines. The programs generate artifacts as a byproduct of their operation, which can be used as evidence. While these artifacts might be essential or non-essential for the system’s functionality (Freiling et al., 2015), the diversity and instability are “dramatically increasing the requirements and complexity of data exploitation tools” (Garfinkel, 2010).<sup>1</sup> So, though practically relevant, building an encyclopedia of digital traces, as Gross and Geerds (1977) or Kirk (1974) started for physical evidence, seems only partly meaningful from a research perspective as long as the fundamental mechanism of the structured translation of investigative questions to relevant traces is not well understood yet. Given the fast-paced technological advances, the present article addresses this issue by taking a step back to take a different route and explore a model-based approach to abstractly de-

---

<sup>1</sup>Sometimes such developments even have implications on the acquisition process such that so-called “app-downgrading” is needed to gather evidence (Geus et al., 2023).

---

\*Corresponding author

Email address: felix.freiling@fau.de (F. Freiling)

ORCID(s): 0000-0003-1862-2900 (J. Gruber); 0000-0002-8279-8401 (F. Freiling)

scribe a structured solution to this problem.

## 1.2. Contributions

To approach this task and contribute to the—still relatively sparse—body of the understanding of cybercriminalistics, we combine our previous works, i.e., Gruber et al. (2022), Gruber and Humml (2023), and Gruber et al. (2023b), aiming to close the evident gap in the translation process from hypotheses to evidence. To the best of our knowledge, the article shows a novel way how to find relevant digital evidence given some investigative questions in a structured way. We aggregate and synthesize previous insights into the Cyber-traceological Model, which can both be regarded (1) as a concrete method to “compute” relevant evidence given favorable circumstances, and (2) as a general thinking model for providing a structured solution to the cybercriminalistic task.

We claim that up to now, there has been no precise and succinct understanding to derive a plan of action for this translation of hypotheses to relevant digital evidence. The Cyber-traceological Model offers such an understanding. So even if the proposal may appear to be hardly directly applicable, we describe the core mechanics of the cybercriminalistic task, which allows both researchers and practitioners to take a new view of existing processes. We argue that by doing so, we can transfer model-based thinking used in computer science into the discipline of cybercriminalistics and emphasize criminalistic understanding in digital forensic science. In addition to that we provide a clear and novel distinction between digital forensic science from cybercriminalistics.

## 1.3. Outline

The remainder of the article is structured as follows: In Section 2, we provide an overview of the necessary background material and refer to related work. We introduce terms like cybercriminalistics, the (cyber)criminalistic task, and traceology, besides essential concepts such as necessary and sufficient evidence as well as the relevance of traces. Then, we develop the Cyber-traceological Model as the core contribution of this work in Section 3. We describe the considerations of a model-based approach, the model’s components, and their interplay. In Section 4, we illustrate the usefulness of our proposal by providing two examples. Afterward, we discuss the merits and the limitations of the model as well as the general potential of such an approach in Section 5 before we give an outlook on future work in Section 6 and conclude the article in Section 7.

## 2. Background and Related Work

*Understanding Traceology & Cybercriminalistics.* Digital forensics, digital investigations, and cybercriminalistics are not clearly delineated—sometimes these terms are even used interchangeably in common parlance; for the overarching field of forensic science, Ristenbatt III et al. (2022) recently sought terminological clarity. To avoid conflation of the terms “forensic science” and “criminalistics”, they

proposed to use the term “traceology” as a more precise version of describing the holistic study of traces. This term goes back to the 1920s, when early criminalists at the Humboldt-Universität Berlin coined the locution “traceology”<sup>2</sup> (Margot, 2011, p. 97). By reviving and using this phrase, Margot (2011, 2014) sparked that development to increasingly focus on the trace and its study aiming to reconstruct past events—aptly named “traceology”.

**Definition 2.1** (Traceology according to Ristenbatt III et al. (2022, p. 29)). *Traceology* is the “[s]tudy of event traces created during an event, which encompasses the detection, recognition, identification, process of individualization toward source attribution, and evaluation of the physical record created (be it an item, pattern, or signal) [...]”.

Traceology can be considered to be one part of the wide-ranging profession and research discipline called criminalistics, when we employ the broader understanding of one of its founders, Hans Gross (1977). According to his notion, traceology is just one component besides the phenomenological aspects of crime commission, criminalistic tactics, and organizational and strategical concerns. Shifting our gaze to the digital domain, we can state that while digital forensic science, as defined by Palmer (2001), is terminologically well thought out, we can see that it can be further divided into two areas: its core deemed *forensic computing* and the more applied topic of *digital investigations*. The former deals with the assessment of associations while the latter is concerned with the entire process of handling and processing digital evidence for forensic purposes, including but not limited to recovery as well as inspection of and search for evidence and its management (Dewald and Freiling, 2014, p. 6); however, cybercriminalistics has not yet been included in this scheme.

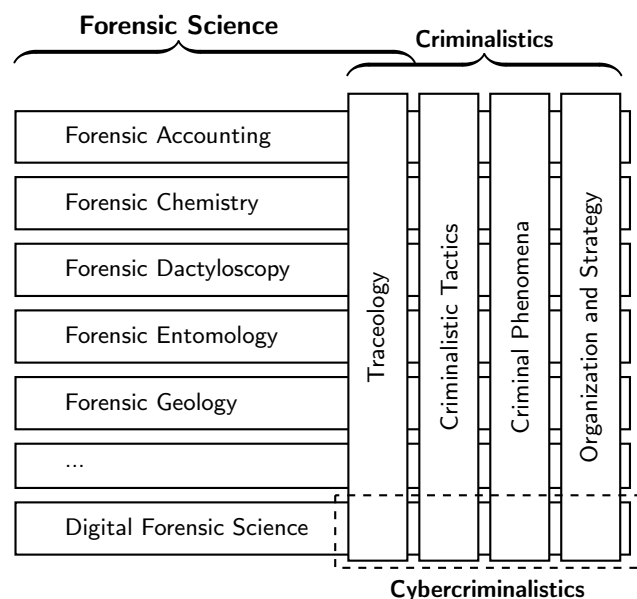
In view of the comprehensive interpretation of the term “digital forensic science”, there arises a need to demarcate it from “cybercriminalistics”—two terms that are unfortunately often used interchangeably, although there are arguably big differences in their meanings. Reminiscing Hans Gross’ understanding of general criminalistics, we recognize an expansion of the scope of the general terms by the incorporation of the digital domain. It becomes apparent that the newly emerged digital dimension of several aspects of criminalistics can be grouped and subsumed by the term “cybercriminalistics”. Hence, we propose to define the field of cybercriminalistics as follows:

**Definition 2.2** (Cybercriminalistics). *Cybercriminalistics* is the digital dimension of the profession and scientific discipline of combatting crime. It comprises the study of the organization and strategy of cybercrime fighting and specifically digital investigations, the study of cybercriminal phenomena and their investigation using traceology of digital traces and digital criminalistic tactics.

In essence, we consider grouping the single branches of forensic science as traceology, i.e., the holistic study of

<sup>2</sup>The original term in German is “Spurenkunde”.

traces. Traceology is then considered part of criminalistics, which combines tactics, phenomenology, organization, and strategy. Cybercriminalistics then comprises the digital dimension of these fields and is primarily concerned with digital traces, digital investigations, and combatting cybercrime, as illustrated in Fig. 1.



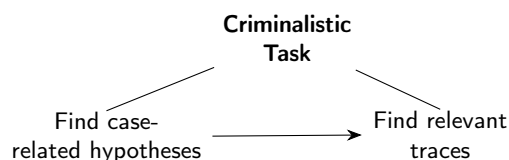
**Figure 1:** The proposed view of the relation of the intertwined disciplines of forensic science, criminalistics, and cybercriminalistics. Referring to Definitions 2.1 and 2.2, we consider grouping the single branches of forensic science as traceology, i.e., the holistic study of traces. Criminalistics builds upon traceology and combines tactics, phenomenology, organization, and strategy, while it is pervaded by the insights of forensic science. Cybercriminalistics then comprises the digital dimension of these fields and is primarily concerned with digital traces, digital investigations, and combatting cybercrime. Referring to this visualization, we want to stress that the present publication's focus is cyber-traceology, i.e., the left-most part of the dashed rectangle.

*The (Cyber)Criminalistic Task.* Confronted with some offense, criminalists need to discover the truthful course of actions retroactively. According to Walder and Hansjakob (2016), the criminalistic task is identified by the following three main quests:

1. Identify crime,
2. gather relevant evidence in flawless manner, and
3. critically assess the acquired evidence to check whether the perpetrator can be convicted or not.

To detail these main tasks, they specify several more detailed questions that need to be answered in every case. These revolve around hypotheses concerning objective and subjective facts, perpetration, unlawfulness and guilt as well as the circumstances relevant to sentencing (Walder and Hansjakob, 2016, p. 6). If we boil the criminalistic task down

to its very essence, we can say that there are two related but distinct subproblems: The first one is the quest to put up hypotheses of pertinence for the investigation, and the second one is to find relevant traces that can be used to assess those previously identified case-relevant hypotheses, as illustrated in Fig. 2. This figure simplifies the problem since the both subtasks are interdependent and the found traces might also allow the assessment of other (not yet explicitly formed) hypotheses.



**Figure 2:** We propose to divide the criminalistic task into two subproblems, i.e., the formation of apt hypotheses and the search for traces relevant for the previously identified hypotheses, which is a difficult quest in the digital domain due to the volume and complexity of the data.

Turning the head to the cyber-dimension and defining the *cybercriminalistic task*, we can restrict the general notion of the criminalistic task to cybercrimes, which does not change much but reduces the breadth of offense and focuses the digital investigative toolset.

*Relevance of Traces.* In view of the criminalistic task set out above, we previously approached the research question of when digital evidence is considered to be relevant and expressive yet reliable (Gruber and Humml, 2023). We derived formal notions for the concepts of *relevance* and *expressiveness of perceivable facets* of digital tangible traces (Jaquet-Chiffelle and Casey, 2021) concerning investigative hypotheses—essential attributes that remained implicit in the field of digital forensic science but provided valuable insights regarding their meaningfulness for an investigation. Building on this clarification, we constructed an investigative knowledge base rooted in formal understanding.

**Definition 2.3** (Investigative knowledge base according to Gruber and Humml (2023)). An *investigative knowledge base*  $(H, F, \text{supports}, \text{refutes})$  consists of a set  $H$  of hypotheses where each element provides a possible explanation of the facets, a set  $F$  of perceivable facets of the traces potentially present at crime scenes where each element is a digital object on a deliberate abstraction level together with two relations,  $\text{supports} \subseteq F \times H$  and  $\text{refutes} \subseteq F \times H$ , relating facets to hypotheses with the expected meanings.

It is worth noting that these two relations need to be disjoint, i.e.,  $\text{supports} \cap \text{refutes} = \emptyset$ , in order to consider the knowledge base to be consistent. Furthermore, there is no logical connection between these two relations. That is, one cannot infer that a facet that is not refuting a hypothesis is necessarily supporting it (and vice versa); still, it might

be consistent with the hypothesis, but it cannot be deemed relevant then.

In this regard, the formalization allows to precisely express such relevance as a relation  $\text{relevant} \subseteq F \times H$ , so that a facet  $f$  relevant  $h := f$  supports  $h \cup f$  refutes  $h$ . Briefly expressed informally in natural language, we say that a facet is relevant to an investigative hypothesis if it either supports or refutes it. Given a hypothesis  $h \in H$  and a set of facets  $F$ , we denote by  $F|_h$  the set of facets in  $F$  that relate to  $h$ :

$$F|_h := \{f \in F \mid f \text{ relevant } h\}$$

The expressiveness  $H|_f$  of a facet in regard to a set of hypotheses, in turn, is characterized by the wealth of hypotheses that a facet can assess. Devising these concepts allowed us to make the so-called criminalistic cycle, a process model to solve the previously described criminalistic task, more precise by introducing exact termination and facet identification—named the Facet-oriented Criminalistic Cycle.

*Necessary and Sufficient Evidence.* Gladyshev and Patel (2004) and Carrier and Spafford (2006) independently developed the idea of viewing digital forensic event reconstruction as a problem to be solved within a (finite state) formal model of the digital system. While the aim of Carrier and Spafford (2006) was to systematize and integrate different practical approaches of digital forensic analysis, Gladyshev and Patel (2004) were the first to devise methods that could *compute* answers to forensic questions. The basic idea is to apply *backtracing* within the formal system model to answer questions about the computational past of a system—an idea that was further developed in subsequent research (James et al., 2009; Soltani and Hosseini-Seno, 2019; Dewald, 2015).

Inspired by these formalization efforts, we considered the more general questions of the usefulness of different classes of evidence in forensic event reconstruction (Gruber et al., 2023b). This work generalizes the quest for finding concrete evidence in a case into a characterization of two different classes of relevant evidence on a technical level: evidence that is either necessary or sufficient for concluding the occurrence of an event in the past. This improves Dewald’s characteristic evidence method (Dewald, 2015) to solve a task that is known as the specific reconstruction problem (SRP). Aiming to define the attributes of necessity and sufficiency of digital traces in forensic event reconstruction in a more complete sense, we employed an automata-theoretic approach and used linear-time temporal logic to define these two fundamental evidence classes: Sufficient evidence (SE), i.e.,  $(\bigcirc\sigma) \mathcal{R}(\neg SE)$  expressed in linear-time temporal logic (LTL), allows to prove the execution of the target action  $\sigma$  in the automaton by containing all the facets that can only be observable if the action has been executed. Necessary evidence (NE), i.e.,  $\bigcirc(\square(\sigma \rightarrow \square NE))$  expressed in LTL, allows to refute the execution of the target action  $\sigma$  by stating all the facets that must be observable in all subsequent states after the execution of  $\sigma$ . Using model checking software, these notions of general reconstructability classes

can be practically used to calculate evidence sets of the evidence classes named. In Section 3, we will show how these concept can be used to materialize relevance and expressiveness as a basis to construct an investigative knowledge base. However, since the use of this approach is confined to rather simple system models, we now turn our heads to a more real-world-oriented approach.

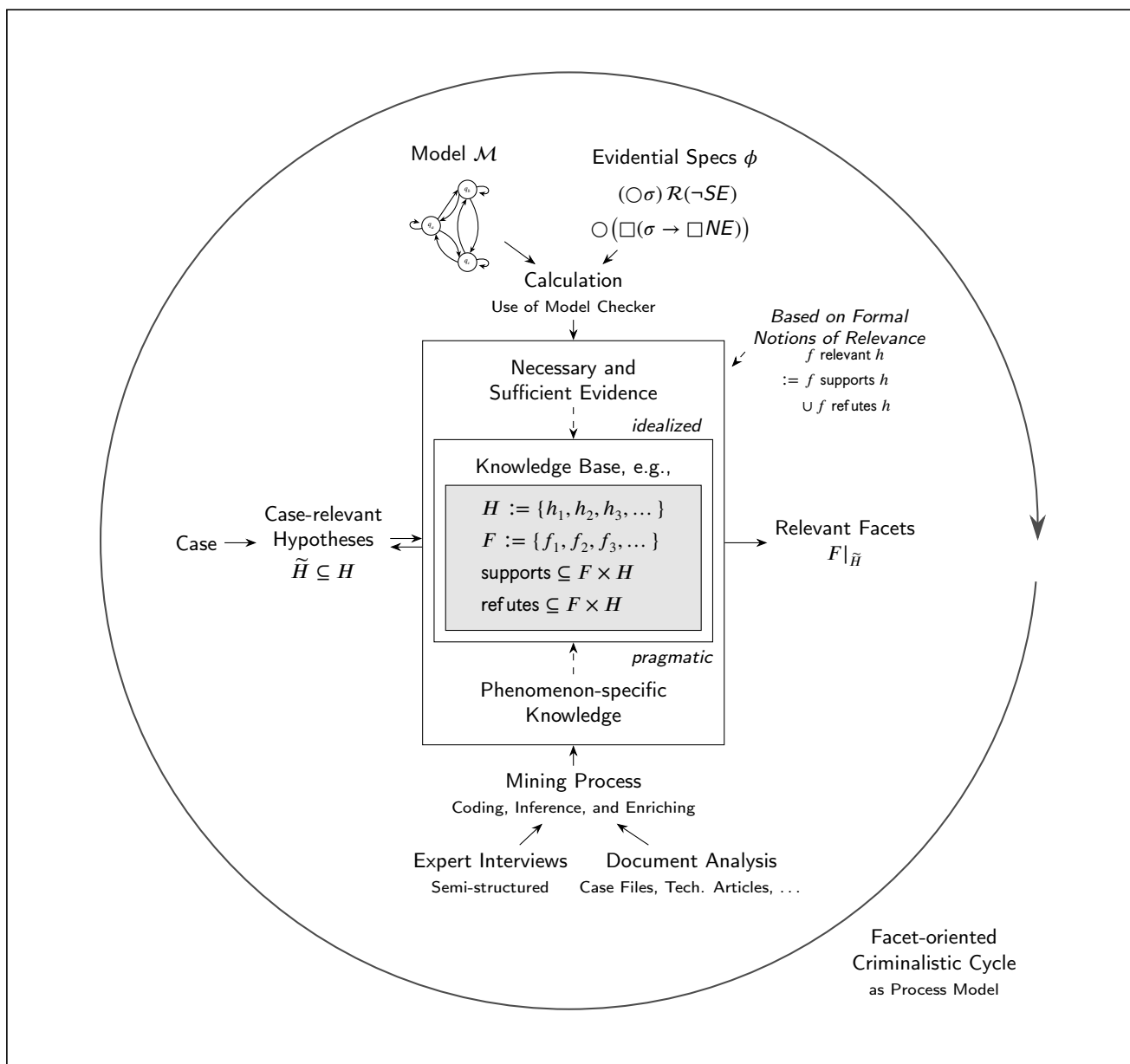
*Phenomenon-specific Knowledge.* From a more practical point of view, we previously looked at the identification of relevant evidence and identified that there is an abstraction gap between universal process models and the concrete proceedings in a specific case (Gruber et al., 2022). To bridge this gap, we proposed the use of phenomenon-specific knowledge, which has been demonstrated by the example of botnet crime. While this is a specific instance, there is still need for a more generic way of guiding investigations on the meso-level of abstraction. We addressed the research question of how telling evidence can be found and documented for use in real-world investigations to achieve criminalistic goals. To introduce such an urgently required intermediary step bridging the abstraction gap, we proposed the collection and use of phenomenon-specific knowledge to encode what is relevant evidence when investigating a specific criminal phenomenon. Using phenomenon-specific investigative knowledge bases in the form of cognitive maps provides practitioners with clear guidance. This approach introduces a generalized yet actionable description of evidence, capturing specific knowledge about the phenomenon and thus supporting the quest to find relevant traces in a more pragmatic setting. To demonstrate the applicability of this method, we presented an exemplary cognitive map of technical investigations in cases of botnet crime, the correctness and completeness of which has been validated by conducting interviews with domain experts.

In the course of the paraphrases presented above, we hinted gently at a common thread running through these individual insights, which might not be apparent when reading the works in isolation. All of these results can be considered components of a structured solution to the criminalistic task. The thoughtful combination of these insights and their integration into the bigger picture of the criminalistic task led to the development of the Cyber-traceological Model.

### 3. The Cyber-traceological Model

#### 3.1. Opting for a Model-based Approach

While most forensic disciplines operate on concrete, tangible evidence, the nature of traces in digital forensic science is more abstract. Computer scientists in general are used to work in intangible environments; hence, they commonly simplify structures for two reasons: Firstly, they impose restrictions on the models to prevent any ambiguity or confusion, as marginal phenomena can be excluded by deliberately limiting the model. Secondly, simplifying the model to focus on the essential components, using symbolic notations, and removing any unnecessary details can help to gain



**Figure 3:** The Cyber-traceological Model depicts the overall process and the basic building blocks to provide a structured method of translating investigative hypotheses to relevant facets by employing an investigative knowledge base that can be built using the different approaches previously proposed by us (Gruber et al. (2022), Gruber and Humml (2023), and Gruber et al. (2023b)).

insights into complex connections that are hard to grasp otherwise. Restrictions of the matter may lead to an expansion of the understanding; vagueness is eliminated on the one hand, and implicitness is turned outward on the other, which is a benefit that outweighs the potential risk of introducing precision and limiting its meaningfulness, as discussed later. So, we are confident to use these characteristics to facilitate reasoning about the cybercriminalistic task.

### 3.2. Description of the Model

The *Cyber-traceological Model* provides a structured method of generically translating investigative demands to the relevant traces, as illustrated in Fig. 3. We argue that it houses the essential building blocks of the cybercriminalistic

task, such that we cannot strip out any of the core components without losing a substantial element and the ability to solve cases. It is important to note however, the model draws from the articles summarized in Section 2 and aggregates their insights to form a universal view of the essentials of criminalistic reasoning.

*The Components.* We place an investigative knowledge base, as defined in Definition 2.3, at the core of the Cyber-traceological Model—and at the figure’s center. It keeps track of the facets and the hypotheses of which the investigators are aware of. Furthermore, it contains the relations that map facets to hypotheses to denote that they either support or refute them. Hence, this core component manifests

the formal notion of relevance, as exemplarily shown by the definition of the relevance relation as a union of the supports and refutes relations in the upper right-hand side of Fig. 3.

The major question is, of course, how the investigative knowledge base can be constructed and correctly filled: In an idealized setting, where a model  $\mathcal{M}$  of the system under investigation is available, the investigative knowledge base can be filled using the evidential specifications, as put up in our previous work (Gruber et al., 2023b), to calculate sufficient and necessary evidence sets for the actions available in the automata with a model checker, as illustrated at the top of Fig. 3. This is possible because of the inherent connection of the concepts of NE/SE and the notion of relevance, i.e., the former mapping hypotheses to facets and the latter facets to hypotheses (Gruber and Humml, 2023). This duality allows a transformation by looking at which evidence sets related to which action the facet occurs so that we can map facet in question to hypotheses it supports or refutes.

Obviously, the ability to employ the NE/SE approach to fill the investigative knowledge base requires the ability to model the system under investigation accurately, which would provide objective rigour but is unfortunately not possible for reasonably complex software yet. A more pragmatic (though rather subjective) method to construct the investigative knowledge base geared toward real-world application is the mining of phenomenon-specific knowledge, as previously proposed by us (Gruber et al., 2022), which is depicted at the bottom of Fig. 3. Here, we proposed a method to collect phenomenon-specific knowledge, which can be used to map artifacts, connected investigative measures, and their potential results in form of node-link knowledge representations. Our method, which employs a mining process, that is constituted by coding, inference, and enriching, allows collecting relevant and expressive facets regarding a specific criminal phenomenon based on experiential knowledge from different sources, such as documents or domain experts.

*Their Interplay.* If the knowledge base has been filled, either constructed by employing the automata-theoretic approach, by using phenomenon-specific knowledge, or any other conceivable method, it can be applied to support investigators in their casework in the following way:

When they are confronted with a case, they need to solve the criminalistic task, as sketched out in Section 2. So, they gather suspicion regarding one or multiple possible criminal offenses and accordingly either come up with hypotheses  $\tilde{H}$  on the objective and subjective facts, which are a subset of  $H$ , which stored in the investigative knowledge base. Note that there is an intersection between the construction of case-related hypotheses and their answering, as indicated by the opposing arrowheads, since the knowledge base contains helpful or crucial information to form apt hypotheses. Then, the investigators query the knowledge base with the case-related hypotheses  $\tilde{H}$  to derive relevant facets that can be used to assess these, as shown on the right-hand side of Fig. 3. Hence, the knowledge base provides a conception of which facets to collect in order to assess the set of case-

related hypotheses by bringing the set of relevant facets  $F|\tilde{H}$  to light. Those facets will help to assess (at least) the hypotheses in  $\tilde{H}$  according to the knowledge base and, hence, can be used as a basis for collecting certain facets. Since it can be expected that some facets might not be collectible or those being collected might refute some hypotheses, there needs to be an iterative process, in which the translation is embedded; thus, we propose using the Facet-oriented Criminalistic Cycle (FoCC) (Gruber and Humml, 2023, Fig. 2) as an enclosing process model, which surrounds the figure circularly.

In the first step, the available or collected facets are assessed by looking for the respective facet in the supports and refutes relations. Then, the case-relevant hypotheses are updated. The iterative updates of the set of case-related hypotheses  $\tilde{H}$  will in turn lead to continually collecting facets until the investigator determines the investigation to be decisively or exhaustively complete. Decisive completeness is reached if all hypotheses in  $\tilde{H}$  can be decided by the collected facets in conjunction with the relations stored in the investigative knowledge base. Exhaustive completeness is assumed if the maximum amount of facets relevant to the investigation, i.e.  $F|\tilde{H}$  as a whole, has been considered—regardless if they have either been successfully collected or could not be recovered. As long as this property is not achieved, the unanswered hypotheses have to be determined in the fourth step, which is the basis for querying the investigative knowledge base for the missing facets, which have to be collected in the final step, as it has been defined as part of the FoCC. This provides an apt procedural frame for structured translation, which completes the Cyber-traceological Model.

#### 4. Examples

To illustrate the usefulness of the Cyber-traceological Model as shown in Fig. 3, we present two examples. By doing so, we aim to demonstrate the adequacy of the concept and try to explain and contextualize what the Cyber-traceological Model is about. The first example deals with a solely theoretical setting to demonstrate the foundational nature of the model and the general feasibility of its construction in an idealized setting, where we have a full-fledged system model. The second one demonstrates how the model can be used in actual investigative work to answer commonly posed investigative questions in a case of possessing and distributing child sexual abuse material (CSAM) as one branch of computer-enabled crimes. There, it is shown that systematic procedures to identify traces helping with the assessment of an investigative hypothesis are still applicable when no formalization is available. These two example provide complementary viewpoints on the same matter: The first exemplification illustrates a formal approach where a model of a fully specified system is available, whose properties we can calculate in a well-defined manner. The second one reflects a practical viewpoint and shows the projection of the Cyber-traceological Model in a real-world scenario, where no full specification of the system is available.

#### 4.1. Interpretation of a System Model's State

At first, we describe the practical incorporation of our previously proposed NE/SE method (Gruber et al., 2023b) into the Cyber-traceological Model.

The basis of this example is a system model  $\mathcal{M}$  in the form of a labeled transition system, in which the states are identified by a set of atomic facts that are true in a particular state and transitions that denote the possible actions in the system leading to state changes. We can form an investigative knowledge base  $KB_{\mathcal{M}}$  specific to a transition system, as laid out in Definition 2.3, by taking the model to derive various attributes: First, we retrieve the labels attached to the transitions  $\sigma_i \in \Sigma$  to come up with the various actions, which are present in  $\mathcal{M}$ . The hypotheses on their execution forms then the set  $H_{\Sigma}$  as the first building block of  $KB_{\mathcal{M}}$ . Second, we can extract the model's variables  $V$  and their possible valuations. These constitute the facets, which are devised by partial valuations of  $V$ .<sup>3</sup>

Having constructed the sets  $F$  and  $H_{\Sigma}$ , we can start to build the supports and refutes relations. To do so, the proof-of-concept implementation calculates the evidence sets,  $NE(\sigma_i, \mathcal{M})$  or  $SE(\sigma_i, \mathcal{M})$  respectively, where a specific action  $\sigma_i$  is mapped to facets, which are sufficient or necessary evidence for its execution. Building upon the previously identified duality of the concepts of NE/SE and facets' relevance (Gruber and Humml, 2023), we can create the inverse mapping to transform this representation into the relations  $\text{supports} \subseteq F \times H$  and  $\text{refutes} \subseteq F \times H$ , relating facets to hypotheses with the expected meanings. For example, such an inverse mapping may be calculated using a naive algorithm, which iterates over all  $f_i \in F$  and checks the respective facet's presence in the various evidence sets to find these actions for which it is evidential allowing to generate an element in the respective relation. Finally, the formation of these two relations then completes the investigative knowledge base  $KB_{\mathcal{M}}$  specific to the system model in question. Having  $KB_{\mathcal{M}}$  at hand, the investigators can now assess the expressiveness of a given facet  $H_{\Sigma}|_f$  or query all relevant facets for a given hypothesis  $F|_h$  to check the encountered system state.

Aiming to demonstrate the concrete realization, we implemented this process, as shown in Fig. 4.<sup>4</sup> To practically do so, we specify the models using the input specification language of the established model checker NuSMV and leverage the Python module formerly provided alongside with our previously mentioned publication introducing the NE/SE method (Gruber et al., 2023b).<sup>5</sup> In essence, our proof-of-concept implementation shows the principal applicability of the Cyber-traceological Model, the connection of the concepts, and the powerful idea of having means to query a knowledge base for relevant facets when a system model is available.

<sup>3</sup>We consider *partial valuations* to be “formulae of the form  $a = v$  where  $a$  is a variable in  $V$  and  $v$  is a value in the range of  $a$ , with every variable mentioned at most once” (Gruber et al., 2023b).

<sup>4</sup>See <https://github.com/jgru/investigative-knowledge-base>.

<sup>5</sup>See <https://github.com/jgru/evidential-calculator/>.

#### 4.2. Interpretation of P2P-Software Traces

After the preceding example, where a system model is required, we now describe a real-world application dealing with the investigation of the distribution of CSAM via peer-to-peer (P2P) networks, which is still the predominant pathway for spreading this kind of material (Steel et al., 2023). Since this niche of computer-enabled crimes constitutes both a widespread and a well-researched area of digital investigations (Adelstein and Joyce, 2007; Liberatore et al., 2010b,a; Hurley et al., 2013), we consider this to be an apt field to refer to for the sake of illustrating the application of our proposed model in a real-world context. To do so, we envision the following scenario:

Via the so-called CyberTipline an electronic service provider reported the download of alleged CSAM to the National Center for Missing and Exploited Children,<sup>6</sup> which forwards the information to the competent law enforcement agency. Referring to stock data of the user account in question, a suspect is identified and a search warrant is obtained from the competent judge because of the suspicion of the possession of prohibited material. So, investigators search the premises of the suspect and seize a computer as evidence item #001.

To clear the case, the investigators need to assess several hypotheses:

$h_1$ : CSAM is present on the storage media of evidence item #001?

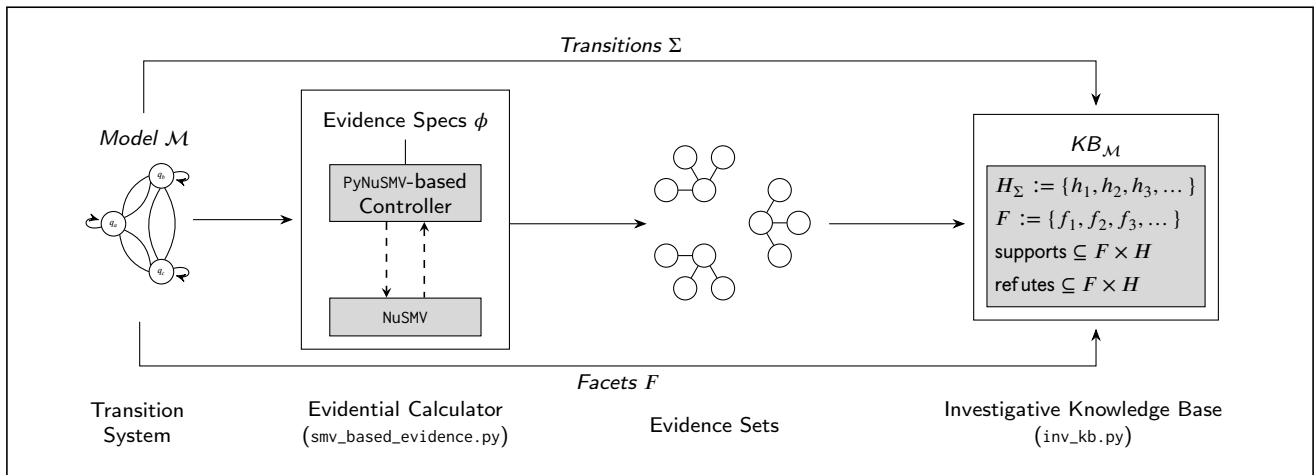
$h_2$ : CSAM has been distributed via P2P networks using evidence item #001?

For the assessment of  $h_1$ , we query the investigative knowledge base and it outputs a set of cryptographic and/or perceptual hashes<sup>7</sup> of known incriminated files as relevant traces  $F|_{h_1}$ . Each hash can be considered a facet in the supports relation and is hence deemed relevant for assessing  $h_1$ . The presence of these files, i.e., elements in  $F|_{h_1}$ , is then checked by traversing all allocated and unallocated files on the storage media of the suspect; if there is at least one match,  $h_1$  can be considered to be correct. Of course, if no match has been found the assessment is not complete since—most certainly—incriminated files are circulating that are not yet known by the law enforcement agencies (LEAs); thus, a content inspection would be necessary, which can be seen as an endeavor to extend the supports relation of the knowledge base.

To assess  $h_2$ , we need a multi-step approach. First, we check the hypothesis  $h_{2,1}$ , i.e., “P2P software is present on the computer system”. Querying the investigative knowledge base, the investigators retrieve a list of relevant traces  $F|_{h_{2,1}}$ , such as hashes of P2P executables, prefetch

<sup>6</sup>See <https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

<sup>7</sup>See <https://www.phash.org>



**Figure 4:** Overview of the construction of an investigative knowledge base as described in Section 4.1. Based on a transition system provided in NuSMV’s input specification syntax, the variables and actions in the automaton are extracted first. The variables’ partial valuations form the facets and the proposition of action executions form the hypotheses. Second, evidence sets are calculated. These are then inverted to map the facets to action executions and build the supports and refutes relations, which complete the investigative knowledge base.

files related to P2P programs, and file paths of configuration files. Checking their presence, the investigators might be able to verify  $h_{2,1}$  and, as a side-effect, identify the respective P2P software used to refine  $h_2$ . For the sake of the example, we assume the presence of the common software *eMule* (Klier et al., 2023); so, the investigators need to pose and assess  $h_{2,2}$ , i.e., “*eMule* has been used to distribute CSAM”. They query the investigative knowledge base for  $h_{2,2}$  and retrieve  $F|_{h_{2,2}}$ . This set of relevant facets contains the filepath to the so-called *known.met* file, which provides—besides several other usage statistics—information about how many bytes of a certain file have been uploaded into the P2P network.  $h_{2,2}$  can then be further refined to  $h_{2,2,1}$ , i.e., “a file classified as CSAM is listed as uploaded in the *known.met* file”. By checking each file path listed as an upload entry in the *known.met* file, the investigators can assess  $h_{2,2,1}$  and might conclude that at least one file constituting alleged CSAM has been seeded.

Klier et al. (2023) recently underlined the practical relevance of these scenarios, where the tracking of file usage and distribution is key for investigations of such kind. While they focused on the “gap between the investigators’ needs, i.e., the automatic extraction of a file-trace, and the capabilities of common forensic applications”, we stress that it firstly needs the notions developed here to substantiate the reasoning for implementing tool support and secondly provide a clear understanding for explainability of investigative conclusions as well as expert witness work.

Of course, this rather simple investigation could also be solved drawing from experiential knowledge of experienced investigators; however, in the era of cybercriminalistics with a digital, extensively connected, and data-driven world, we consider IT supported processes to be vastly helpful, if not essential, to effectively handle the wealth of diverse and complex traces.

## 5. Discussion

After having described the Cyber-traceological Model’s use, we want to confer about its significance, anticipated benefits, and its limitations. Afterward, we broaden our view to ponder about the general potentials of model-based approaches in (digital) forensic science.

### 5.1. Significance and Benefits of the Model

Traceology, as a holistic combination of the branches of forensic science (Margot, 2017; Ristenbatt III et al., 2022), is primarily concerned with assessing hypotheses related to past events based on facets, the observable parts of tangible traces. Recently, the trace and the formal study of its nature have been of increased interest in the forensic science community—aiming to unify traditional and digital branches Jaquet-Chiffelle and Casey (2021). Given the absence of any straightforward method to find “sufficient digital evidence”, the present article consequently took up this development on an abstract and foundational level. The model-based approach aims to solidify reasoning in cyber-traceology and to extend its formal study, viz., the elaboration of the meaning of digital traces for investigative hypotheses. The deduced foundational understanding aims to contribute to solving the second subproblem of the cyber-criminalistic task, i.e., determining relevant digital traces that can be used to assess previously identified case-relevant hypotheses. For doing so, it is salient to have an unambiguous understanding of the attributes of relevance, expressiveness, necessity, and sufficiency of evidence and the interplay of findings, as developed and reflected in the Cyber-traceological Model.

As indicated in the previous section, it becomes obvious that—although we call it the *Cyber-traceological Model* and focus on digital evidence—it is not necessarily specific to the digital domain. The need to translate investigative de-

mands to relevant traces seems to be more general. Maybe one could even reach so far and name the task of finding relevant traces—digital or physical—based on the investigative hypotheses the “holy grail of criminalistics”. Much in the tradition of computer science, we argue that model-based systematics offer help to focus on the absolute essentials, providing an unobstructed view of the matter.

## 5.2. Limitations of the Model

The reduction of the matter, which constitutes a strong point on the one hand, can be considered a downside on the other hand; hence, one can argue that there is a certain imprecision inherent in the investigative knowledge base, which has been placed at the model’s center. This is aggravated by the currently used formalism, as previously proposed by us (Gruber and Humml, 2023), which uses crisp logic and, thus, does not allow to represent probabilities. While this could be solved by resorting to a probabilistic logic approach so that the supports and refutes relations do not merely map facets to hypotheses but provide a probability from the unit interval, another downside is the representation of facets, which is rather vague. Certainly, this is not an issue for portraying the foundations of the model and its applicability in idealized, theoretical settings, it might lead to difficulties in real-world applications, where relevant facets have to be transformed into an actionable representation. In addition, we need to consider the investigative knowledge encoded in the model to be incomplete, much like it is the case for non-structured representations as they are used nowadays; still, the Cyber-traceological Model (and also the encapsulating FoCC) can be updated any time in the process. Turning away from facets, we see that we can identify investigative hypotheses based on a given set of facets by looking into the relations; however, we have to admit that the complexity of real-world hypothesis generation is not fully depicted in the model. Furthermore, the hierarchical representation of hypotheses, as first described by Cook et al. (1998a), and their potential inter-dependences are not represented.

## 5.3. Potentials of Model-based Approaches in Forensic Science

Model-based approaches offer the possibility to structure and streamline processes, as shown by Cook et al. (1998b) with their “model for case assessment and interpretation” for example. Likewise, if rigorous and formal, they allow more foundational insights by creating a simplified representation of a complex system or process. By doing so, researchers may discover that there are key factors that they were previously unaware of, that are not well understood, or that require further investigation. Therefore, this characteristic can help to identify gaps in knowledge and highlight areas where further research is needed. Additionally, such models can be used to communicate complex ideas and concepts in a more concise yet accessible way. Furthermore, models can be used to facilitate the development and testing of new theories and ideas. By creating a simplified representation of a complex system or process, researchers can look at it in a fo-

cused way, test different hypotheses, and explore the potential outcomes of different scenarios. This can help to identify new patterns and relationships, leading to the development of new theories and ideas.

By using model-driven approaches, we suspect that they have the potential to provide universal insights and—with some fantasy—one might even imagine that the theories developed by digital forensics could rise and become metatheories for other disciplines of the field. Such a development would then constitute a unification of the disciplines, which is a development that has, for example, already been gently sparked by our previous work (Gruber et al., 2023a), in which we proposed a common definition of evidence contamination; as computer scientists, we argue that there seems to be tentative potential that other branches of forensic science could profit from the results of reasoning about the abstract nature of things using models, as it is natural and maybe even imposed when working with digital systems.

## 6. Future Work

The new insights gained by the fusion of previous work have been manifested in the creation of the Cyber-traceological Model. Having scrutinized the model, we see both foundational and practical threads of future work.

*Practical Tasks.* Having a solidified notion of investigative knowledge bases and their use for solving the criminalistic task, we envision an ongoing collection of phenomenon-specific knowledge for all significant cybercrime phenomena, e.g., various types of online fraud (i.e., investment fraud, romance scams, and others), ransomware, CSAM, dark web narcotics trafficking, to build up an encyclopedia of facets and the hypotheses assessed by them for real-world application—much like the vision of influential researchers of modern-day criminalistics, such as Gross and Geerds (1977) and Kirk (1974).

The quest to build up encyclopedias, however, exacerbates the actuality of the previously mentioned question of how to represent facets. For this purpose, we need further research to explore the most effective (and usable) representations of facets in investigative knowledge bases. This also raises the question of whether to build facet descriptions on top of an ontological model. Since forensic artifacts have numerous been described in such a way (Barnum, 2012; Harichandran et al., 2016; Syed et al., 2016; Casey et al., 2017), we intend to investigate the application of such an approach. A promising candidate seems to be the use of the *CASE* ontology proposed by Casey et al. (2017), which extends the *Unified Cyber Ontology* proposed by Syed et al. (2016), to represent facets. While this ontological approach is used in real-world software such as *Hansken* developed and implemented by the Netherlands Forensic Institute (van Beek et al., 2015), the general applicability for physical traces has to be investigated.

Promisingly, the use of an ontological model could build the basis to explore automated reasoning in the future. It can be suspected that such an aid will enhance both the overview

and situational awareness of the case under investigation. However, even without the rather far-fetched idea of automated reasoning, we need to measure the efficacy of the Cyber-traceological Model in supporting investigations. To do so, we envision collaborations with practitioners to conduct user studies and assess the impact of task-relevant information in the form of phenomenon-specific knowledge bases on the analysis results, much like Sunde and Dror (2021) did for investigating “biasability” of examiners by providing task-irrelevant information.

**Foundational Questions.** We identified two limitations revolving around hypotheses: First, there is a need for modeling and representing probabilities, as already identified in one of our previous works (Gruber and Humml, 2023). Second, the procedure to come up with an investigative hypotheses, i.e., the first part of the criminalistic task, is not well understood yet.

While reporting the value of digital evidence in the form of probabilities (or more precisely, likelihood ratios) is still uncommon, there seems to be a need to quantify probabilities linked to investigative questions for digital evidence; this, however, requires further studies to complement the initial forays by Kwan et al. (2008), Tse et al. (2012), and Overill and Silomon (2010). Considering the criminalistic task, as visualized in Fig. 2, we see that the Cyber-traceological Model helps to solve the second subproblem, i.e., the search for relevant traces. We directly see the potential to extend the model dealing with uncertainty by using probabilistic semantics and swap the supports and refutes relations for single-valued functions projecting facets and hypotheses to a rational value on the unit interval encoding the likelihoods, which will—in principle—allow us then to resort to a likelihood ratio-based approach.

Regarding the second issue, we see that there are many open questions that revolve around the quest to find methods that empower the investigator to systematically come up with apt investigative propositions, i.e., finding those hypotheses exhibiting case-related relevance, and to refine as well as relate those to one another. Existing work in that regard, i.e., the hierarchy of propositions by Cook et al. (1998a) only dealt with the general classification of hypotheses. The real issue lies in the structured and hierarchical generation of investigative hypotheses, which requires (experimental) knowledge of specific criminal phenomena (Brodag, 2001, p. 303 f.) and even a great deal of creativity, experience, and mental openness to develop apt hypotheses, facilitating both intuitive and reflexive thinking methods in combination (Walder and Hansjakob, 2016, pp. 173 ff.). Unsurprisingly, this is a quest that can be considered a vastly complex problem in the field of argumentation theory (Bex, 2021); hence, we suggest a deep exploration of the mechanisms of hypothesis formation in future work.

## 7. Conclusion

Solving crimes is an age-old endeavor. The recent increase of pervasive computing in all areas of human life in-

creases the relevance of a whole new discipline called cyber-criminalistics and also uncovers issues in criminalistic thinking in general and traceology in specific that have been dormant for centuries. Thinking about the criminalistic task, we see that the complexity and quantity problems of digital data exacerbate the question of what constitutes “relevant digital evidence”. To tackle this demand, we survey our previous works (Gruber et al. (2022), Gruber and Humml (2023), and Gruber et al. (2023b)) to consolidate these insights in a novel model-based approach to support solving the cybercriminalistic task. For doing so, it is salient to have an unambiguous understanding of the attributes of relevance, expressiveness, necessity, and sufficiency of evidence and the interplay of findings, as developed in this article and reflected in the Cyber-traceological Model. The proposed model provides both conceptual clarity and prospective practical guidance in identifying relevant traces in the vast sea of digital data. Interestingly, the tackled questions are, in their more profound nature, similar to what pioneers like Edmond Locard (1920) or Hans Gross (1977) had faced back in time. However, their convolution seems to be amplified by the features of the digital domain, especially by quantity and complexity problems of evidence in fast-changing IT environments (Carrier, 2003). By facing these, the present article aims to work at the foundation and contribute to the understanding of fundamental connections and attributes linked to digital evidence employing a model-based approach; still, the article also pointed out many contemporary questions, more or less fundamental, for future research directions.

## Acknowledgements

We thank Lena Voigt and Christian Lindenmeier for their comments on an early draft. Work has been supported by DFG (German Research Foundation) as part of the Research and Training Group 2475 “Cybercrime and Forensic Computing” (grant number 393541319/GRK2475/2-2024).

## CRedit authorship contribution statement

**Jan Gruber:** Conceptualization, Methodology, Investigation, Writing - Original draft, Writing - Review & Editing, Visualization. **Felix Freiling:** Conceptualization, Funding Acquisition, Methodology, Supervision, Writing - Review & Editing.

## References

- Adelstein, F., Joyce, R.A., 2007. File marshal: Automatic extraction of peer-to-peer data. *Digit. Investig.* 4, 43–48. URL: <https://doi.org/10.1016/j.diin.2007.06.016>, doi:10.1016/J.DIIN.2007.06.016.
- Anderson, R.J., Barton, C., Böhme, R., Clayton, R., van Eeten, M.J.G., Levi, M., Moore, T., Savage, S., 2013. Measuring the cost of cybercrime, in: Böhme, R. (Ed.), *The Economics of Information Security and Privacy*. Springer, pp. 265–300. doi:10.1007/978-3-642-39498-0\_12.
- Barnum, S., 2012. Standardizing cyber threat intelligence information with the structured threat information expression (stix). *Mitre Corporation* 11, 1–22.
- van Beek, H.M.A., van Eijk, E.J., van Baar, R.B., Ugen, M., Bodde, J.N.C., Siemeling, A.J., 2015. Digital forensics as a service: Game on. *Digit.*

- Investig. 15, 20–38. URL: <https://doi.org/10.1016/j.diin.2015.07.004>, doi:10.1016/J.DIIN.2015.07.004.
- Berghel, H., 2017. On the problem of (cyber) attribution. *Computer* 50, 84–89. URL: <https://doi.org/10.1109/MC.2017.74>, doi:10.1109/MC.2017.74.
- Bex, F., 2021. Argumentation and evidence. *Philosophical Foundations of Evidence Law*, 183–198. doi:10.1093/oso/9780198859307.003.0014.
- Brodag, W.D., 2001. *Kriminalistik – Grundlagen der Verbrechensbekämpfung. Kriminalistik und Kriminologie*. 8 ed., Richard Boorberg Verlag, Stuttgart, Germany.
- Carrier, B.D., 2003. Defining digital forensic examination and analysis tool using abstraction layers. *Int. J. Digit. Evid.* 1.
- Carrier, B.D., Spafford, E.H., 2006. Categories of digital investigation analysis techniques based on the computer history model. *Digit. Investig.* 3, 121–130. URL: <https://doi.org/10.1016/j.diin.2006.06.011>, doi:10.1016/j.diin.2006.06.011.
- Casey, E., Barnum, S., Griffith, R., Snyder, J., van Beek, H.M.A., Nelson, A.J., 2017. Advancing coordinated cyber-investigations and tool interoperability using a community developed specification language. *Digit. Investig.* 22, 14–45. doi:10.1016/j.diin.2017.08.002.
- Cook, R., Evett, I., Jackson, G., Jones, P., Lambert, J., 1998a. A hierarchy of propositions: deciding which level to address in casework. *Science & Justice* 38, 231–239. URL: <https://www.sciencedirect.com/science/article/pii/S1355030698721173>, doi:10.1016/S1355-0306(98)72117-3.
- Cook, R., Evett, I., Jackson, G., Jones, P., Lambert, J., 1998b. A model for case assessment and interpretation. *Science & Justice* 38, 151–156. doi:10.1016/S1355-0306(98)72099-4.
- Dewald, A., 2015. Characteristic evidence, counter evidence and reconstruction problems in forensic computing. *Inf. Technol.* 57, 339–346.
- Dewald, A., Freiling, F., 2014. *From Computer Forensics to Forensic Computing: Investigators Investigate, Scientists Associate*. Technical Report CS-2014-04. Friedrich-Alexander-University Erlangen-Nuremberg (FAU). URL: [https://opus4.kobv.de/opus4-fau/files/4750/computer\\_forensics\\_is\\_not\\_forensic\\_science.pdf](https://opus4.kobv.de/opus4-fau/files/4750/computer_forensics_is_not_forensic_science.pdf).
- Freiling, F.C., Schuhr, J.C., Gruhn, M., 2015. What is essential data in digital forensic analysis? *Inf. Technol.* 57, 376–383. URL: <http://www.degruyter.com/view/j/itit.2015.57.issue-6/itit-2015-0016/itit-2015-0016.xml>.
- Garfinkel, S.L., 2010. Digital forensics research: The next 10 years. *Digit. Investig.* 7, S64–S73. URL: <https://doi.org/10.1016/j.diin.2010.05.009>, doi:10.1016/J.DIIN.2010.05.009.
- Geus, J., Ottmann, J., Freiling, F., 2023. Systematic evaluation of forensic data acquisition using smartphone local backup, in: *Proceedings of the Digital Forensics Research Conference Europe (DFRWS USA)*, dfrws.org. pp. 1–13. URL: <https://dfrws.org/presentation/systematic-evaluation-of-forensic-data-acquisition-using-smartphone-local-backup/>.
- Gladyshev, P., Patel, A., 2004. Finite state machine approach to digital event reconstruction. *Digit. Investig.* 1, 130–149. doi:10.1016/j.diin.2004.03.001.
- Gross, H., Geerds, F., 1977. *Handbuch der Kriminalistik: Wissenschaft und Praxis des Verbrechensbekämpfung*. M. Pawlak.
- Gruber, J., Hargreaves, C.J., Freiling, F.C., 2023a. Contamination of digital evidence: Understanding an underexposed risk. *Forensic Sci. Int. Digit. Investig.* 44, 301501. doi:10.1016/j.fsidi.2023.301501.
- Gruber, J., Humml, M., 2023. A formal treatment of expressiveness and relevance of digital evidence. *Digital Threats* doi:10.1145/3608485.
- Gruber, J., Humml, M., Schröder, L., Freiling, F.C., 2023b. Formal verification of necessary and sufficient evidence in forensic event reconstruction, in: *Bajramovic, E., Rodríguez, R.J. (Eds.), Proceedings of the Digital Forensics Research Conference Europe (DFRWS EU)*, dfrws.org, Bonn. pp. 1–11.
- Gruber, J., Voigt, L.L., Benenson, Z., Freiling, F.C., 2022. Foundations of cybercriminalistics: From general process models to case-specific concretizations in cybercrime investigations. *Forensic Sci. Int. Digit. Investig.* 43, 301438. doi:10.1016/J.FSIDI.2022.301438.
- Harichandran, V.S., Walnycky, D., Baggili, I.M., Breiting, F., 2016. CuFA: A more formal definition for digital forensic artifacts. *Digit. Investig.* 18 Supplement, S125–S137. URL: <https://doi.org/10.1016/j.diin.2016.04.005>, doi:10.1016/J.DIIN.2016.04.005.
- Hurley, R., Prusty, S., Soroush, H., Walls, R.J., Albrecht, J., Cecchet, E., Levine, B.N., Liberatore, M., Lynn, B., Wolak, J., 2013. Measurement and analysis of child pornography trafficking on p2p networks, in: *Proceedings of the 22nd International Conference on World Wide Web*, Association for Computing Machinery, New York, NY, USA. p. 631–642. URL: <https://doi.org/10.1145/2488388.2488444>, doi:10.1145/2488388.2488444.
- James, J., Gladyshev, P., Abdullah, M.T., Zhu, Y., 2009. Analysis of evidence using formal event reconstruction, in: *Goel, S. (Ed.), Digital Forensics and Cyber Crime - First International ICST Conference, ICDF2C 2009*, Albany, NY, USA, September 30–October 2, 2009, Revised Selected Papers, Springer. pp. 85–98. doi:10.1007/978-3-642-11534-9\_9.
- Jaquet-Chiffelle, D.O., Casey, E., 2021. A formalized model of the trace. *Forensic Science International* 327, 110941. doi:10.1016/j.forsciint.2021.110941.
- Kirk, P.L., 1974. *Crime investigation*. 2 ed., John Wiley & Sons, Nashville, TN.
- Klier, S., Varenkamp, J., Baier, H., 2023. Back and forth—on automatic exposure of origin and dissemination of files on windows. *Digital Threats* 4. URL: <https://doi.org/10.1145/3609232>, doi:10.1145/3609232.
- Kwan, M.Y.K., Chow, K., Law, F.Y.W., Lai, P.K.Y., 2008. Reasoning about evidence using bayesian networks, in: *Ray, I., Sheno, S. (Eds.), Advances in Digital Forensics IV, Fourth Annual IFIP WG 11.9 Conference on Digital Forensics*, Kyoto University, Kyoto, Japan, January 28–30, 2008, Springer. pp. 275–289. doi:10.1007/978-0-387-84927-0\_22.
- Liberatore, M., Erdely, R., Kerle, T., Levine, B.N., Shields, C., 2010a. Forensic investigation of peer-to-peer file sharing networks. *Digit. Investig.* 7, S95–S103. URL: <https://doi.org/10.1016/j.diin.2010.05.012>, doi:10.1016/J.DIIN.2010.05.012.
- Liberatore, M., Levine, B.N., Shields, C., 2010b. Strengthening forensic investigations of child pornography on p2p networks, in: *Proceedings of the 6th International Conference, Association for Computing Machinery*, New York, NY, USA. URL: <https://doi.org/10.1145/1921168.1921193>, doi:10.1145/1921168.1921193.
- Locard, E., 1920. *L'enquête criminelle et les méthodes scientifiques*. Bibliothèque de philosophie scientifique, E. Flammarion.
- Margot, P., 2011. Forensic science on trial—what is the law of the land? *Australian Journal of Forensic Sciences* 43, 89–103.
- Margot, P., 2014. Traçologie: la trace, vecteur fondamental de la police scientifique. *Revue internationale de criminologie et de police technique et scientifique* 67, 72–97.
- Margot, P., 2017. Traceology, the bedrock of forensic science and its associated semantics, in: *Rosy, Q., Décary-Héту, D., Delémont, O., Mulone, M. (Eds.), The Routledge international handbook of forensic intelligence and criminology*. Routledge, pp. 30–39.
- Overill, R.E., Silomon, J.A., 2010. Digital meta-forensics: quantifying the investigation, in: *Proc. 4th International Conference on Cyber-crime Forensics Education & Training (CFET 2010)*, Canterbury, UK (September 2010).
- Palmer, G., 2001. A road map for digital forensic research. Technical Report DTR-T001-01 Final. Air Force Research Laboratory, Rome, New York.
- Ristenbatt III, R.R., Hietpas, J., De Forest, P.R., Margot, P.A., 2022. Traceology, criminalistics, and forensic science. *Journal of Forensic Sciences* 67, 28–32.
- Soltani, S., Hosseini-Seno, S., 2019. A formal model for event reconstruction in digital forensic investigation. *Digit. Investig.* 30, 148–160. doi:10.1016/j.diin.2019.07.006.
- Steel, C.M.S., Newman, E., O'Rourke, S., Quayle, E., 2023. Technical profiles of child sexual exploitation material offenders. *Psychiatry, Psychology and Law* 0, 1–14. doi:10.1080/13218719.2022.2148305.
- Sunde, N., Dror, I.E., 2021. A hierarchy of expert performance (HEP) applied to digital forensics: Reliability and biasability in digital forensics decision making. *Digit. Investig.* 37, 301175. doi:10.1016/j.fsidi.2021.301175.
- Syed, Z., Padia, A., Finin, T., Mathews, M.L., Joshi, A., 2016. UCO: A unified cybersecurity ontology, in: *Martinez, D.R., Streilein, W.W., Carter,*

- K.M., Sinha, A. (Eds.), Artificial Intelligence for Cyber Security, Papers from the 2016 AAAI Workshop, Phoenix, Arizona, USA, February 12, 2016, AAAI Press. URL: <http://www.aaai.org/ocs/index.php/WS/AAAIW16/paper/view/12574>.
- Tse, H., Chow, K., Kwan, M.Y.K., 2012. Reasoning about evidence using bayesian networks, in: Peterson, G.L., Sheno, S. (Eds.), Advances in Digital Forensics VIII - 8th IFIP WG 11.9 International Conference on Digital Forensics, Pretoria, South Africa, January 3-5, 2012, Revised Selected Papers, Springer. pp. 99–113. doi:10.1007/978-3-642-33962-2\_7.
- Walder, H., Hansjakob, T., 2016. Kriminalistisches Denken. 10 ed., Kriminalistik, C.F. Müller.