NOTIONS OF HYPOTHESIS IN DIGITAL FORENSICS

Segen Tewelde, Stefan Gruner, Martin Olivier

HAL Id: hal-01449069

https://hal.inria.fr/hal-01449069

Submitted on 30 Jan 2017

Chapter 2

NOTIONS OF HYPOTHESIS
IN DIGITAL FORENSICS

Segen Tewelde, Stefan Gruner and Martin Olivier

Abstract   With the growing scientification of the discipline of digital forensics, the notion of "scientific hypothesis" is becoming increasingly important, because all empirical science is hypothetical, not apodictic. Although the word "hypothesis" is used widely in the digital forensics literature, its usage is not sufficiently reflected from a philosophy of science point of view. This chapter discusses this problem with particular reference to Carrier's methodological work in which the notion of hypothesis plays a prominent role.

Keywords: Digital forensics, philosophy of science, scientific hypotheses

1.   Motivation

This chapter is based on the little-disputed premise that the young discipline of digital forensics is still – to use a phrase by Thomas Kuhn [16] – in a "proto-scientific" stage, although its scientification is on the way [11]. When a discipline grows from a proto-scientific stage to a "paradigmatic" stage [14], two phenomena can be typically observed: (i) a growing formalization (or mathematization) in the expression of the theories in the discipline; and (ii) a growing methodological awareness of what is required or forbidden or allowed for the methods in the discipline in order to qualify as "scientific." These phenomena manifest themselves according to Bunge's differentiation between the substantive and operative parts [6] of the entire body of knowledge of a discipline.

As far as the first (substantive) point is concerned, the mathematical formalization of the theoretical language in digital forensics has already shown some progress (see, e.g., [4, 8, 20]). On the one hand, this is because scholars from the traditional theoretical computer science and formal methods communities are successfully penetrating the interdisci-

plinary barriers and are beginning to gain visibility in the field of digital forensics. On the other hand, the need for formalization has also been recognized by a number of scholars in the digital forensics community. One example is the formal description of aspects of digital forensics by Cohen [9], including critiques of earlier projects.

The second (operative) point is more problematic in the development of digital forensics. Not only is there little methodological awareness, but the growing methodological disputes in the discipline are often taking place outside the realm of science. For example, according to Cohen [9], many of the details of currently-used procedures and tools are neither published nor thoroughly tested. Where tool testing is discussed, the coverage is often reminiscent of testing in an engineering sense (e.g., a tool performs some task correctly) as opposed to an assessment in a methodological sense (i.e., the task performed ought to be at the methodological core of the discipline).

Carrier [7] has contributed to this discourse by formulating a hypotheses-oriented approach for all digital forensic investigations to meet the classical science-philosophical criteria of scientificness. (Note that the term "science-philosophical" is used throughout this chapter as the adjective form of the philosophy of science.) Alas, not every hypothesis is *per se* scientific. Consequently, the scientificness of an empirical discipline such as digital forensics is not automatically guaranteed by a demonstration that it follows (or is able to follow) a hypothesis-based approach. Additional scientificness criteria, such as those identified by Bunge [5], must be stipulated for hypotheses before they can be used as indicators of the degree of scientificness of the discipline in which the hypotheses are formulated.

At this point, one of the methodological weaknesses in the current meta-theory of digital forensics has been reached. Whereas the word "hypothesis" is already in wide use (see, e.g., [2]), it is mainly used in an undifferentiated and science-philosophically unreflected manner. This suggests that users of the concept are largely unaware of the many different classes of empirical hypotheses. A potentially harmful illusion of scientificness could arise as a consequence of this ignorance.

The foregoing considerations motivate this meta-analysis of Carrier's methodological framework [7] with regard to Bunge's well-established classification scheme for scientific hypotheses [5]. This is, indeed, the principal contribution of this chapter. Note that this chapter is not concerned with the often-mentioned extrinsic pragmatic value of the scientificness of digital forensics (i.e., acceptability in courts of law). Instead, this chapter focuses on the intrinsic philosophical value of scientificness

for its own sake, whereby the notion of scientificness must never be conflated with truth.

2.   Related Work

The use of Bayesian decision networks for evaluating and interpreting scientific findings in forensic science is discussed in [12, 18]. The philosophical problems underling the narrative construction of explanations by means of hypothetical linkages between individual events are detailed in [10]. Note that the work of Carrier [7] also has a strong focus on the explanatory construction of such kinds of histories [17] of (digital) events.

An appropriate meaning of the term "hypothesis" is crucial for all these considerations because genuine science must be an explanatory undertaking [22], which goes well beyond the mere gathering and description of data and facts (as the descriptivists would have it). Thereby, all empirical/scientific explanations are ultimately hypothetical [6].

The formulation of hypotheses for explanatory purposes is also discussed in the related field of forensic engineering [21]. Thereby, the types of hypotheses formulated in digital forensics are similar to those formulated in software testing because both digital forensics and software testing are idiographically oriented towards what is unique (not towards what is universally general). For this reason, the meta-scientific methodological considerations about the problem of justification of empirical hypotheses in software testing [1] should also be studied carefully by digital forensic theoreticians.

3.   Carrier's Work

Carrier [7] has proposed a methodological framework for hypothetical digital forensic event reconstruction with high scientific plausibility. The framework incorporates seven categories and more than thirty classes of analysis techniques for formulating or testing digital forensic hypotheses. Before going into the details of Carrier's categories, it is instructive to use an example [7] to understand the conceptual relationships between the categories and the purposes for which the hypotheses in the categories are formulated.

> **Example:** A computer is suspected of having been used to download contraband pictures from the Internet. Lower-level hypotheses are formulated about the time that the system was operational and the technical (operational) capabilities of the system.
>
> The hypotheses are used as presuppositions for formulating succeeding hypotheses of higher complexity. At this point, the lower-level presuppositions are tacitly considered to be true.

Next, a higher-level hypothesis is formulated that asserts disjunctively that either program A or program B was used to download the pictures from the Internet. This hypothesis is formulated on the basis of the assumption that programs A and B are capable of downloading the pictures.

In order to determine which of the two programs was actually applied, a particular log file ℓ is inspected to see which program was operational during the suspected time of the incident. From a logical point of view, the existence of the log file ℓ justifies the deduction:

$$\frac{(A \vee B),\ \neg B}{A}$$

under the condition $\ell \implies \neg B$. This investigative step, however, hypothetically assumes the trustworthiness (integrity) of the log file itself because otherwise the required implication $\ell \implies \neg B$ would no longer be compelling.

Last, but not least, it must be taken into account that the computer could have been compromised by an intruder who stored the pictures on the computer without the computer owner's knowledge.

Given the need to create different types of hypotheses with various degrees of plausibility, Carrier [17] has proposed a schema for formulating hypotheses about (possible) sequences of past events in the lifetime of digital evidence. The remainder of this section examines the schema in detail.

3.1   History Duration

According to Carrier, hypotheses in this category are formulated about the time span T during which the system under analysis was operational. To formulate testable hypotheses in the History Duration category, lower-level hypotheses about time instants $t \in T$ must be formulated.

With regard to Bunge's classification scheme [5], the hypotheses are typically singular and specifiable because they can be obtained by substituting variables for constants to account for single facts. They are also testable with regard to the existence of log files or other temporal traces. As far as Bunge's precision attribute is concerned, hypotheses in the History Duration category are refined as they are both predicate-precise and range-precise. However, they are typically isolated and not systemic as they define unique events in a manner that Windelband [22] called idiographic as opposed to nomothetic. Hypotheses of this kind are typically, in Bunge's terms, both confirmable as well as refutable in principle. Note, however, that confirmability and refutability in principle cannot guarantee *de facto* confirmation or refutation in every individual case.

Carrier notes that hypotheses about an entire duration T are not without exceptions. For example, all the information on a storage device can be overwritten and, thus, traces of previous events can be erased [7]. From a science-philosophical point of view, this corresponds to the situation of historians and archaeologists such that the science-philosophical methodologies of the two classical disciplines can become relevant (and important) to digital forensics [17]. Hypothetical claims about the previous existence of information that no longer exists are typically unspecifiable (in Bunge's terms), but they can motivate deeper research in their role as programmatic hypotheses.

3.2 Primitive Storage System Configuration

According to Carrier, the first type of hypotheses in this class have the form: device d has a storage capacity c, which is empirically unknown before a measurement is taken. A hypothesis of this type is, again, singular and specifiable in Bunge's terminology. The hypothesis is also phenomenological because it merely describes the surface of a phenomenon, not the details of its inner mechanisms. The hypothesis is both confirmable and refutable by means of a suitable experimental apparatus. Interested readers are referred to [6] for a deep science-philosophical discussion of observations, measurements and experiments.

The second type of hypotheses in the Primitive Storage System Configuration category have the form: device d was connected during the time T of an incident. In Bunge's terms, this is again a singular hypothesis of localizing existential character. Because it refers to the past, the previously-mentioned historical/methodological issues [17] are also relevant to this class, especially where the possibility of deleted information or information that no longer exists are concerned.

3.3 Primitive Event System Configuration

Carrier's third category covers hypotheses that make assertions about the capabilities of devices involved in events of interest. In Bunge's classification scheme, these correspond to the same types of hypotheses discussed above.

3.4 Primitive State and Event Definition

In this category, Carrier normatively states that five classes of techniques shall be used to formulate and test hypotheses: (i) primitive state observation class; (ii) state and event capability class; (iii) state and event sample data class; (iv) state and event reconstruction class; and (v) state and event construction class.

Activities in the primitive state observation class are meant to collect data by observing an output device like a computer monitor. The observed state is defined in the inferred history for the times that the observation was made. The corresponding hypothesis deals with a single state that existed at a specific time. In Bunge's terms, this is a singular, specifiable, testable, refined and grounded hypothesis that is experience-referent and phenomenological. Carrier states that the data should be considered as facts if they do not conflict with the observations made by another investigator. Thus, Bunge's considerations about observations, facts and phenomena in the intersection of the knowing subjects and the physical objects [6] are especially relevant.

In Carrier's state and event capability class, hypotheses about primitive system capabilities correspond to presuppositions. They must be tested and accepted before they can be used to formulate hypotheses about possible system states and events. This composition of higher-level and lower-level hypotheses is, in Bunge's terms, mechanismic because of the references to the inner workings of the machinery under test. Once again, the hypotheses are singular, specifiable, refined and grounded.

Hypotheses formulated with techniques in the state and event sample data class are probability hypotheses that cannot be decisively confirmed or refuted by finite amounts of data. Similar types of hypotheses occur (to use one of Bunge's examples) in the field of clinical medicine where tobacco smoking and lung cancer are typically correlated with each other. Carrier has clarified that the research techniques choose events and states that occur with a probability above some threshold value. Some level of subjectivity must be admitted because neither the threshold value nor the notion of probability are clearly specified. What may appear as highly probable to one investigator may be improbable to another investigator.

Carrier's requirements for the state and event reconstruction class are that an investigator must understand the logic associated with the event capabilities of a system and that it would be prudent to identify the unique signatures of events (this provides background information for formulating a hypothesis). A hypothesis formulated by this technique has the form: event e causes state s. In Bunge's classification scheme, this is the strongest type of hypothesis because it must be defined functionally in terms of one free stimulus variable and its dependent effect variable, and it must be thoroughly experimented with in order to distinguish genuine causality from mere coincidence. At this point, Carrier may well consider the possibility of generating general nomothetic hypotheses of the form: every event of type e will always cause a

state of type s. This may be incorporated in a case-independent general theory of digital forensics.

For the final technique, the hypotheses formulated in the state and event construction class are probability hypotheses too, with their notorious difficulties as far as confirmability and refutability are concerned. The truth value of such a hypothesis can only be assessed based on the end state of a trace. However, quasi-general hypotheses [5] can be formulated in cases where a system prevents, by its own construction, certain events from occurring when the system is in a particular state.

3.5   Complex Storage System Configuration

Two types of hypotheses are formulated in this category, which defines entities such as the names of complex storage types that existed during a time interval T, the transformation functions for each of the complex storage types, the attribute names for each of the complex storage types, etc. The hypotheses in this category are more complicated than hypotheses in the simpler History Duration category. Here, the techniques defined for the hypotheses do not test them decisively because probability hypotheses are at best confirmable, not refutable. The hypotheses typically state that a particular device d existed during a time interval T. Since the complex storage was created by a program, a forensic investigator needs to identify the corresponding program. For this purpose, Carrier specified a program identification technique that searches for reconstructed program states. The previously-mentioned science-philosophical problems of historiography, especially with regard to securely deleted information, are also relevant [17].

The second type of hypotheses in this category, the complex storage capability hypotheses, are singular hypotheses about the capabilities of devices. Carrier proposed three techniques to formulate and test these hypotheses. Since all three techniques are related to testing software, it important to consider the science-philosophical issues pertaining to the justification of empirical hypotheses in software testing discussed by Angius [1].

3.6   Complex Event System Configuration

The first of the three types of hypotheses formulated in this sixth category is basically the same as in the Complex Storage System Configuration category described above; its associated techniques have also been analyzed above. Using the data type reconstruction class of techniques in this category, the digital forensic investigator formulates hypotheses that are difficult to test. Once again, these hypotheses are similar

to those formulated in scientific software testing [1]; the objects of the hypotheses are algorithms and data structures. Because algorithms and data structures have many-to-many relations with each other – one algorithm can manipulate many data structures and one data structure can be manipulated by many algorithms – the notorious Duhem-Quine dilemma [19] looms for the hypotheses. From the perspective of Bunge's classification scheme, these are hypotheses of relatedness.

3.7   Complex State and Event Definition

Carrier identifies only one type of hypothesis in the final Complex State and Event Definition category. Eight classes of techniques are applicable to support hypothesis formulation: (i) state and event system capability class; (ii) state and event sample data class; (iii) state and event reconstruction class; (iv) state and event construction class; (v) data abstraction class; (vi) data materialization class; (vii) event abstraction class; and (viii) event materialization class.

A technique in the first class, based on the complex system capabilities, is used by an investigator to determine a possible state or event from the list of possible states and events associated with the system of interest; this bears much similarity with the primitive state and event system capability class. Here, the investigator formulates the same type of hypotheses, with the only difference that, for these hypotheses, information about system capabilities is presupposed from prior work in the previous categories.

In the state and event sample data class, a standardized sample is required for all investigations if an objective hypothesis is to be formulated. Again, this does not differ very much from the primitive state and event sample data class discussed above. In Bunge's classification scheme, this corresponds (once again) to a singular, specifiable, testable, refined and grounded hypothesis.

Techniques in the state and event reconstruction class refer to the occurrence times of events in multitasking distributed systems. Empirical hypotheses in such contexts often have a statistical character because of the notorious non-deterministic behavior of distributed systems in which certain input/output observations are often not experimentally reproducible.

Techniques in the state and event construction class can be used to postulate events that may have occurred, albeit with a rather low level of confidence on the part of the investigator. Bunge has categorized such vague hypotheses in his pragmatic functions category [5].

Techniques in the data abstraction class generate hypotheses by checking the data in the inferred history to determine the complex storage location types that could be relevant. These are singular and testable hypotheses that can be refuted by comparing the range of each attribute to the attribute value defined by the previous complex state event system capability technique. The Duhem-Quine dilemma again looms when the data in the inferred history can support multiple complex storage types.

Hypotheses produced using the data materialization class are similar to those discussed above.

Unlike the scenarios involving complex storage locations, lower-level events can only be part of one complex event for a specific inferred history. Grounded and testable hypotheses can be formulated in such scenarios unless there are multiple lower-level events, in which case, a new inferred history is created. Thus, an inferred history represents an investigator's assumptions about the events in an incident. If inconsistencies arise, an alternative inferred history must be identified. For the given inferred history, however, singular, specifiable and grounded hypotheses are formulated using the descriptions of lower-level events. A technique in the event abstraction class is used to refute a hypothesis about the complex events when lower-level events for the same time are defined in the inferred history, but are not caused by the event that is the focus of the hypothesis.

The techniques belonging to the final event materialization class are not used to formulate hypotheses. Instead, they are used to refute primitive or complex event hypotheses in cases where a higher-level event existed at the same time T, but was not causally related to the lower-level event stated in the (refuted) hypothesis.

4. Bunge's Classification

Bunge's classification scheme for scientific hypotheses [5], which has been used in the previous section as the science-philosophical basis of the present analysis of Carrier's work, is extremely thorough, fine-grained and subtle. Many of Bunge's categories are especially relevant to the – in Windelband's terms – nomothetic [22] sciences, in which hypotheses serve the highest purpose in formulating general scientific laws (e.g., Einstein's $E = mc^2$ in theoretical physics) or quasi-general statistical laws or hypotheses (e.g., smoking causes cancer in the domain of medical science).

In digital forensics, the situations in which hypotheses are formulated are usually fundamentally different in that the general epistemic interest of digital forensics is typically not nomothetic. Digital forensic profes-

sionals are not interested in proposing general law-like hypotheses such as $E = mc^2$. Instead, they are primarily interested in finding out what has happened and when and why in historically unique and hardly generalizable situations. In Windelband's terms [22], digital forensics is thus idiographic rather than nomothetic.

As far as Bunge's classification scheme is concerned, it is not surprising that the specific and individual fact-referent classes of hypotheses occur most frequently in Carrier's framework. Digital forensics shares these idiographic characteristics with the science of software testing, in which (too) only specific hypotheses about a given system under test are formulated [1]. Most relevant to digital forensics is Bunge's classification of hypotheses with regard to their testability. These include [5]:

- Empirically untestable hypotheses

- Purely confirmable hypotheses

- Purely refutable hypotheses

- Both confirmable and refutable hypotheses

Also relevant is Bunge's classification of hypotheses with respect to their pragmatic functions [5]:

- Generalizers of past experiences

- Case-specific inference starters

- Programmatic research guides

- Explanatory hypotheses

- *Ad hoc* protectors of other hypotheses

From this point of view, the degree of scientificness of every discipline increases with its amount of confirmable and refutable explanatory hypotheses.

In this context, however, it must be noted that Bunge emphasized that truth and scientificness are not synonyms: let $\mathcal{H} = \mathcal{H}_s \uplus \mathcal{H}_n$ be a disjoint union of scientific and non-scientific hypotheses, and let $I : \mathcal{H} \to \mathbb{B}$ be a Boolean interpretation of the hypotheses. Then, it is possible for two hypotheses that $I(h) = \mathsf{t}$ and $I(h') = \mathsf{f}$ with $h \in \mathcal{H}_n$ and $h' \in \mathcal{H}_s$. The pragmatic implications to digital forensics are obvious. Indeed, this is a very important, but often forgotten, issue.

Similarly, the manner by which a hypothesis comes into existence is not relevant to the scientificness of the hypothesis. Accordingly, Bunge's scientificness criteria for hypotheses [5] are not genealogical criteria. In

other words, even the most sophisticated scholarly method could generate non-scientific hypotheses and even the proverbial "random monkey" could generate a hypothesis that formally satisfies all the criteria of scientificness.

5.   Limitations of the Study

Carrier's methodological framework is normative rather than descriptive. Carrier provides guidelines to digital forensic practitioners about the types of hypotheses that ought to be formulated, as well as about the types of objects about which the hypotheses ought to be formulated during the course of forensic investigations. Of course, what is actually done "in the field" by digital forensic practitioners is quite a different question, one which this study has not addressed.

Because Carrier's normative methodological work resides in an intermediate layer between the science-philosophical analysis discussed here and digital forensic practice, it is not possible to judge the actual degree of scientificness of the digital forensics discipline by assessing Carrier's meta-work about the discipline from the elevated meta-meta-level as engaged in this study. By analogy, it is also not possible to make a judgment about the discipline of physics – in the way it is carried out by physicists in their laboratories – by critiquing Karl Popper's or Hans Reichenbach's philosophies of physics as "scapegoats." This is because Popper and Reichenbach might have been mistaken in their own interpretations of the discipline of physics in such as way that any anti-Popper or anti-Reichenbach critique no longer strikes the actual physicists working in their laboratories.

This distance from the actual usage of hypotheses in the daily practice of digital forensics is clearly a clear shortcoming of the present study. As mentioned below, the present study could have been strengthened by an analysis of how hypotheses are actually formulated and used by practitioners in the field. This would have provided deeper insights into the actual state of scientificness of the digital forensics discipline. Alas, such a field study would have been too difficult to conduct with the available resources. Therefore, Carrier's work was chosen as a *pars pro toto* substitute.

Thus, only if (or so far as) the field behaves according to Carrier's normative prescriptions is the preceding analysis of Carrier's work by implication also an analysis of the field. If, however, Carrier's notions and the practice of the field are disjoint, then the above meta-meta-study about a meta-study can only be – in Bunge's terms – of program-

matic character to motivate further meta-scientific and methodological research.

6.   Conclusions

Hypotheses that have been tested and strongly corroborated over time eventually develop into scientific theories, which are (semantically) tightly connected networks of hypotheses [5]. The development of a scientific theory, however, is not the goal of digital forensic practitioners who formulate hypotheses specifically about the cases they investigate. Nevertheless, case-specific hypotheses must also, according to Bunge [5], be theoretically embedded (grounded), for example in theories of computer science or computer engineering [13].

As far as the scientificness of the notion of hypothesis is concerned – which is not to be confused with the truth of an individual hypothesis – Carrier's work does not strongly contrast with Bunge's scientificness criteria for formulating hypotheses. By and large – although not philosophically-systematically – Carrier's guidelines have indeed taken into account established criteria such as well-definedness and testability.

At this point, it is important to distinguish – again – between science and practice. To count as scientific, a hypothesis must be, in principle, testable (regardless of whether or not it has actually been tested). For practical use, such as in digital forensics, a hypothesis should actually be tested too (which implies its testability). Thus, when Carrier recommends that digital forensic practitioners should sometimes assume some hypotheses to be true in order to get on with their investigations, he does not leave the territory of scientificness as long as the preliminary assumptions are testable in principle. Whether they are actually true or not is not a matter of their scientificness.

A few minor issues can be found in Carrier's framework because it was developed for practical and not philosophical purposes. An example is the somewhat vague conceptual separation between the notions of data, fact and hypothesis in his primitive storage system configuration category.

Due to Carrier's focus on the discipline of digital forensics, the types of hypotheses mentioned in his guidelines are typically singular (case-specific) and often probabilistic (in Bunge's terms), which clearly indicates the idiographic character (in Windelband's terms) of the discipline. While singular hypotheses are often in danger of being too isolated (or *ad hoc*) with regard to embedding theories (and are, thus, deficient in their explanatory power), probabilistic hypotheses do not only have their notorious testability issues with regard to corroboration or refutation by

single instances of observation, but also depend on an often unclarified (i.e., intuitive) notion of probability. Nomothetic hypotheses of high generality and deep (law-like) explanatory power are not prominent in Carrier's work: this is a feature of the digital forensics discipline and is thus certainly not Carrier's fault.

What Carrier has conducted, in contrast with Bunge, is not a classification of the types of hypotheses that may be formulated during digital forensic investigations. Rather, Carrier has systematically partitioned the domain of digital forensic investigations into sub-domains and describes rather informally the material subjects about which hypotheses have to be formulated within the boundaries of the sub-domains. From a methodological point of view, Carrier's work thus provides the digital forensics community with a domain theory – the importance of which must not be underestimated [3] – instead of a type theory of hypotheses. Thereby, from a type-theoretic point of view, the types of hypotheses in Carrier's domains and sub-domains are more or less the same. Furthermore, whereas the methods that Carrier recommends for generating hypotheses are interesting in practice, they are not relevant to the scientificness of the generated hypotheses because scientificness is not a matter of genealogy.

As far as future work is concerned, additional empirical and theoretical research on the concept of hypothesis in digital forensics needs to be conducted. The empirical research should focus on how hypotheses are actually formulated and applied by practitioners in their field work regardless of Carrier's normative stipulations at his methodological meta-level. The theoretical research should attempt to clarify formally how hypotheses can be chained to construct logically-consistent discursive arguments used in forensic reasoning. Both these research efforts might benefit from the improvement and refinement of the domain-theoretic work [3] initiated by Carrier in [7], including the formal mereological [15] considerations at the basis of domain theory.

Acknowledgement

References

[1] N. Angius, The problem of justification of empirical hypotheses in software testing, *Philosophy and Technology*, vol. 27(3), pp. 423–439, 2014.

[2] H. Beyers, M. Olivier and G. Hancke, Database application schema forensics, *South African Computer Journal*, vol. 55, pp. 1–11, 2014.

[3] D. Bjorner, Domain theory: Practice and theories, discussion of possible research topics, *Proceedings of the Fourth International Colloquium on the Theoretical Aspects of Computing*, pp. 1–17, 2007.

[4] G. Bosman and S. Gruner, Log file analysis with context-free grammars, in *Advances in Database Forensics IX*, G. Peterson and S. Shenoi (Eds.), Springer, Heidelberg, Germany, pp. 145–152, 2013.

[5] M. Bunge, *Philosophy of Science: From Problem to Theory, Volume One*, Transaction Publishers, New Brunswick, New Jersey, 1998.

[6] M. Bunge, *Philosophy of Science: From Explanation to Justification, Volume Two*, Transaction Publishers, New Brunswick, New Jersey, 1998.

[7] B. Carrier, A Hypothesis-Based Approach to Digital Forensic Investigations, CERIAS Technical Report 2006-06, Center for Education and Research in Information Assurance and Security, Purdue University, West Lafayette, Indiana, 2006.

[8] Y. Chabot, A. Bertaux, C. Nicolle and M. Kechadi, A complete formalized knowledge representation model for advanced digital forensics timeline analysis, *Digital Investigation*, vol. 11(S2), pp. S95–S105, 2014.

[9] F. Cohen, *Digital Forensic Evidence Examination*, Fred Cohen and Associates, Livermore, California, 2009.

[10] P. Garbolino, Historical narratives, evidence and explanations, in *Explanation, Prediction and Confirmation, The Philosophy of Science in a European Perspective, Volume 2*, D. Dieks, W. Gonzales, S. Hartmann, T. Uebel and M. Weber (Eds.), Springer, Dordrecht, The Netherlands, pp. 293–303, 2011.

[11] P. Garbolino, The scientification of forensic practice, in *New Challenges to Philosophy of Science, The Philosophy of Science in a European Perspective, Volume 4*, H. Andersen, D. Dieks, W. Gonzales, T. Uebel and G. Wheeler (Eds.), Springer, Dordrecht, The Netherlands, pp. 287–297, 2013.

[12] P. Garbolino and F. Taroni, Evaluation of scientific evidence using Bayesian networks, *Forensic Science International*, vol. 125(2-3), pp. 149–155, 2002.

[13] P. Gladyshev, Formalizing Event Reconstruction in Digital Investigations, Doctoral Dissertation, Department of Computer Science, University College Dublin, Dublin, Ireland, 2004.

[14] T. Kuhn, *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, Illinois, 1962.

[15] E. Luschei, *The Logical Systems of Lesniewski*, North-Holland, Amsterdam, The Netherlands, 1962.

[16] M. Olivier and S. Gruner, On the scientific maturity of digital forensics research, in *Advances in Digital Forensics IX*, G. Peterson and S. Shenoi (Eds.), Springer, Heidelberg, Germany, pp. 33–49, 2013.

[17] M. Pollitt, History, historiography and the hermeneutics of the hard drive, in *Advances in Digital Forensics IX*, G. Peterson and S. Shenoi (Eds.), Springer, Heidelberg, Germany, pp. 3–17, 2013.

[18] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino and C. Aitken, *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science*, John Wiley and Sons, Chichester, United Kingdom, 2014.

[19] J. Vuillemin, On Duhem's and Quine's theses, in *The Philosophy of W.V. Quine*, L. Hahn and P. Schilpp (Eds.), Open Court, Peru, Illinois, pp. 595–618, 1986.

[20] J. Wang, Z. Tang and X. Jin, An OCL-based formal method for cloud forensics, *Advanced Materials Research*, vols. 989-994, pp. 1513–1516, 2014.

[21] J. Wiechel, D. Morr and B. Boggess, Application of the scientific method to the analyses in forensic science with case example, *Proceedings of the International Mechanical Engineering Congress and Exposition*, paper no. IMECE2010-39044, pp. 515–522, 2010.

[22] W. Windelband, History and natural science, *Theory and Psychology*, vol. 8(1), pp. 5–22, 1998.