Proposal for a Theoretical Framework in Digital Forensics

Proposal for a theoretical framework in digital forensics

Federico COSTANTINI[a]
[a] *Università degli Studi di Udine (Italy)*

Abstract. This short paper aims to introduce a theoretical framework in digital forensics based on "Philosophy of Information". After a preliminary clarification of its key concepts, some general issues concerning "Information Quality" are outlined in digital and cloud forensics. At the end, I offer a few remarks on future researches' perspectives.

Keywords. Digital forensics, Cloud forensics, Philosophy of Information, Judicial proceedings, Information quality, Philosophy of law

1. Introduction

The concept of proof is crucial in law, placed among epistemology, philosophy of language and theory of argumentation, and – from a strictly legal perspective – between the substantive laws and those governing judicial procedures. With the "Information Society" the very source of evidence[1] has become "information" in itself. Indeed, digital evidence is challenging contemporary legal thought since it is neither an empirical medium (a physical "thing"), nor a witness' statement (an intangible "word")[2].

In "digital forensics"[3] – the forensic discipline applied to digital evidences – ITCs are not just an analytic tool but, indeed, the subject of investigation. Provided such immateriality, very limited trust can be given to this kind of proof since it is complicated to validate the veracity of the source, the accuracy of the analysis and the integrity of the results.

[1] In this paper, for the sake of brevity the words "proof" and "evidence" are used as synonyms, yet their meaning is dissimilar, furthermore if considered comparatively among different legal systems.

[2] Court allegations based upon digital evidence are often expressed in terms of statistical probability so their meaning cannot be qualified neither as empirical finding, nor as full presumption or legal argument.

[3] Digital forensics has been defined as «*the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation, and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations*» [10]. It can be divided in computer forensics, network forensics, cloud forensics, audio forensics, image & video forensics, mobile forensics, embedded forensics, malware forensics, etc. Scholars classify in different taxonomies activities performed in digital forensics and developed various models [8]. A similar field of study, concerning the discussion of electronic evidence in civil proceedings, is called "electronic discovery".

The most recent development of ICTs is due to *cloud computing*, technologies in which resources are partially or entirely "virtualized"[4]. Cloud forensics[5] is a very elusive concept, as there is neither a hard disk to access, nor a network to analyse, or a data stream to intercept[6].

This short paper aims to suggest a theoretical framework in digital forensics. After a preliminary clarification of key concepts in "Philosophy of Information", I outline some general issues of digital and cloud forensics concerning "Information Quality". At the end, I conclude with a few remarks on perspectives for future researches.

2. "Philosophy of Information", LOAs, MASs and judicial proceeding

In cybernetics[7], "information" can undertake three different ontological statuses: "information *as* reality" (technological information)[8]; "information *about* reality" (natural information)[9]; "information *for* reality" (cultural information) [2][10].

"Philosophy of Information"[4][11] brings further this naturalistic vision, aiming to a synthesis between "reality" and "representation", as well as between "object" and "observer". Indeed, according to such perspective, an "information" is considered within its "Level of Abstraction" (LOA), which defines how analysis is performed and thus specifies the criteria used in the observation [7]. The LOA, in other words, represents the point of view adopted by the observer, namely, it is a formalized model of the observer's expectations concerning analysis' outcomes [5][12]. Furthermore, interaction among many observers can be shaped in a model and represented as a multi-agent-system (MAS)[13].

---

[4] Cloud computing has been defined as «*a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction*» [9]. Depending on how resources are virtualized, three types of cloud services can be provided: IaaS (Infrastructure as a service), PaaS (Platform as a service), SaaS (Software as a service) [9].

[5] Cloud forensics has been defined as «*the application of computer forensics principles and procedures in a cloud computing environment*» [11]. Cloud's contents, by their nature, are extremely volatile and thus relevant data can be changed or no longer be available at the time of trial, due to several factors: for example, voluntary deletion, fortuitous event, data obsolescence, damaged infrastructure, database corruption. Furthermore, and this is the key feature of "virtualization", a cloud does not hold any trace of such changes.

[6] From a forensics' perspective, the key point is that «*evidence can reside everywhere in the world in a virtualization environment*» [14].

[7] As it is well known, it is an interdisciplinary approach aimed at shaping a rationalized unified vision where science, philosophy and spirituality are brought together [1; 17].

[8] For example, the electrical signal, which is transmitted regardless of the message contained.

[9] Such as the information about natural phenomena, which can be true or false (alethic).

[10] In it, instructions or algorithms are conveyed to one or many recipients. These three types of information have been respectively named "technical", "semantic" and "influential" in Weaver's theory of communication [16].

[11] This approach has deepened the ontological, epistemological, and ethical aspects of cybernetics and it is taken into consideration in the "Onlife Manifesto", a document promoted by the European Union where social entailments of such perspective are explained and discussed [6]. See https://ec.europa.eu/digital-agenda/en/onlife-manifesto.

[12] From such perspective, the outcome of an analysis – its "meaning" – requires: (1) the preliminary definition of a LOA; (2) a rigorous epistemic strategy in qualifying the findings as observable "objects". It is possible to set LOAs in which becomes irrelevant the difference not only between *hardware* and *software*, but also between technological protocols and procedural regulations, or even people and machines.

[13] Some scholars argue that each agent in a MAS is charged with an "epistemic responsibility", assuming the duty – qualified almost as an ethical obligation – to gather, organize and share valuable information to enable others making rational decisions and obtain effective results from their interactions [13].

3. Judicial proceedings, evidence and "information quality"

A judicial proceeding can be seen as a LOA set on a different level from the facts to which they refer, since it concerns events that took place formerly and thus need to be represented by parties in order to be discussed and decided[14]. A court trial can be also shaped as a MAS including heterogeneous observers or agents (judges, lawyers, policemen, parties, witnesses, expert witnesses, court clerks, etc.).

We can detect the three kinds of "information" above described also in this context.

Since legal procedures are independent from the substantial cases discussed and constitute the pattern of the LOA, they can be resembled to "information *as* reality". Evidences brought by parties, instead, could be qualified as "information *about* reality" regardless of their nature (written documents or witness hearings) and their empirical appearance (physical supports or electronic data streams). Finally, judge's decision can be qualified as "information *for* reality" as it states what action has to be undertaken as outcome of the trial.

Provided that, forensic sciences have the specific purpose to scrutinize "information" collected from concerned facts, so to guarantee the "information quality" in the trial [12]. Indeed, it is important to remark that forensics' methods and tools need not only to be well-practiced by consultants, but also clearly explained in order to be assessed and discussed even by non-experts (defendants, judges, jurors).

If "transparency" in forensic sciences is a relevant feature of fair trials, it becomes crucial in digital forensics, where evidence is not embodied in a physical entity. Consequently, each phase in the analysis has to allow a complete disclosure in order to be reviewed and discussed[15]. For example, "information" can be extracted by the "forensic image" brought to trial (as in computer or mobile forensics), and "information quality" can be guaranteed by means of the procedures performed – using *open source* software, for example[16] – or providing certified tools for the analysis[17].

In cloud forensics, however, "information quality" is even a more difficult task. Due to the lack of an empirical "observable" entity, the whole acquisition of evidence has to take place entirely in a different LOA and all the process has to be traced in order to fulfil the requirement of transparency[18].

---

[14] According to this view, generally a legal procedure can be defined as a given set of technological processes – natural and artificial, bureaucratic and technical in a strict sense – organized as a workflow ruled by a MAS. It exchanges information with its ecosystem by receiving an *input* – the description of the case to be decided, evaluated or ruled – and by generating an *output* in terms of a legal act (such as a judgment, an administrative act, a regulation).

[15] See for example the "chain of custody" prescribed by the Convention on Cybercrime (Council of Europe), opened to signature on 23 November 2001 in Budapest.

[16] Some scholars claim that there is no need for the software to be completely *open source*, since parties have the possibility to access – and dispute over – the relevant part of source code [15].

[17] In this paper are not discussed issues concerning the use of cloud technology in processing evidence after its collection [3; 15].

[18] Technically, this goal can be achieved building an *ad hoc* "virtualized" environment where relevant data are captured, as provided for example by the LegalEye platform (www.legaleye.it), supported by the Departments of Computer Science and of Legal Science in the University of Udine (Italy).

4. Final remarks

We are witnessing times of strong innovation in all fields of forensics sciences[19].

From a theoretical perspective, I believe it would be very interesting to pursue the following search paths: (1) deepen the representation of the legal procedures in terms of LOA, as seen in "philosophy of information"; (2) define the role of the "quality of information" not only in forensic science (information *about* reality), but also as regards the procedural rules (information *as* reality) and the court decision (information *for* reality); (3) develop a better understanding of cloud forensics; (4) represent in terms of "second order" systems the strategic behaviour of each agent within a legal procedure.

References

[1] G. Bateson, *Mind and nature: a necessary unity*, Dutton, New York, 1979.

[2] A. Borgmann, *Holding on to reality. The nature of information at the turn of the millennium*, University of Chicago Press, Chicago, 1999.

[3] C. Federici, AlmaNebula: A Computer Forensics Framework for the Cloud, *Procedia Computer Science* **19** (2013), 139-146.

[4] L. Floridi, *The philosophy of information*, Oxford University Press, Oxford-New York, 2011.

[5] L. Floridi, *The Ethics of Information*, Oxford University Press, London, 2013.

[6] L. Floridi, ed., *The Onlife Manifesto. Being Human in a Hyperconnected Era*, Springer International Publishing, Cham, 2015.

[7] P. Illari, P. Allo, B. Baumgaertner, S. D'Alfonso, N. Fresco, F. Gobbo, C. Grubaugh, A. Iliadis, E. Kerr, P. Giuseppe, F. Russo, C. Schulz, M. Taddeo, M. Turilli, O. Vakarelov, and H. Zenil, *The Philosophy of Information - a Simple Introduction. Society for the Philosophy of Information*, 2012.

[8] M.D. Kohn, M.M. Eloff, and J.H.P. Eloff, Integrated digital forensic process model, *Computers & Security* **38** (2013), 103-115.

[9] P. Mell and T. Grance, The NIST Definition of Cloud Computing, in, U.S. Department of Commerce, 2011.

[10] G. Palmer, *A Road Map for Digital Forensic Research. Report From the First Digital Forensic Research Workshop (DFRWS)*, New York, 2001.

[11] D. Povar and G. Geethakumari, A Heuristic Model for Performing Digital Forensics in Cloud Computing Environment, in: *Security in Computing and Communications,* J. Mauri, S. Thampi, D. Rawat, and D. Jin, eds., Springer Berlin Heidelberg, 2014, pp. 341-352.

[12] B. Schafer, Information Quality and Evidence Law: A New Role for Social Media, Digital Publishing and Copyright Law?, in: *The Philosophy of Information Quality,* L. Floridi and P. Illari, eds., Springer, Cham; Heidelberg, 2014, pp. 217-238.

[13] J. Simon, Distributed Epistemic Responsibility in a Hyperconnected Era, in: *The Onlife Manifesto,* L. Floridi, ed., Springer International Publishing, 2015, pp. 145-159.

[14] S. Simou, C. Kalloniatis, E. Kavakli, and S. Gritzalis, Cloud Forensics: Identifying the Major Issues and Challenges, in: *Advanced Information Systems Engineering,* M. Jarke, J. Mylopoulos, C. Quix, C. Rolland, Y. Manolopoulos, H. Mouratidis, and J. Horkoff, eds., Springer International Publishing, 2014, pp. 271-284.

[15] H.M.A. van Beek, E.J. van Eijk, R.B. van Baar, M. Ugen, J.N.C. Bodde, and A.J. Siemelink, Digital forensics as a service: Game on, *Digital Investigation* **15** (2015), 20-38.

[16] W. Weaver, The Matemathics of Communication, *Scientific American* **181** (1949), 11-15.

[17] N. Wiener, *Cybernetics or control and communications in the animal and the machine*, Hermann & Cie-The Technology Press, Paris-Cambridge, 1948.

[19] From computer forensics are springing not only mobile forensics, but also drone forensics and robot forensic, as well as in digital forensics great expectations arise in network forensics and, of course, cloud forensics. Further challenges will come from "Internet of Things", where forensics will face the ultimate synthesis of *hardware* and *software*.