Data at the center of the fight against criminality

Éric Freyssinet,
Colonel

Abstract:
The fight against criminality is grounded on the collection, analysis and presentation of data as proof in penal cases. These activities now entail cybercriminalistics, computational and data forensics, and call for exercising command over data.

Thanks to Alphonse Bertillon at the end of the 19th century, collecting data came to be seen as a key to the successful inquests. In the 21st century, the fight against crime and delinquency became centered on collecting, analyzing and presenting data as evidence in penal cases. Fully developing these concrete activities in digital forensics and criminal intelligence involves exercising command over the data.[1]

Cybercriminalistics

Cybercriminalistics, or computational and digital forensics, refers to a set of techniques related to the evidence that, used in legal proceedings, relies on electronic data and digital devices. It should be pointed out in passing that these methods are very similar to those used to undertake investigations in matters related to the security of information systems.

The evolving technology of digital forensics

As techniques and uses have developed, electronic, or digital, evidence (which dates back to the start of the 1980s) is now potentially omnipresent regardless of the offense being investigated. This evidence is usually found on material mediums, but it might also be collected from the electromagnetic environment (*e.g.*, the detection of communications via WiFi in a suspect's home) or even retrieved from third party operators in information and communications technology (ICT), where it has been stored on servers.

What characterizes electronic evidence is its volatility and potential fragility. For example, a computer's operating system constantly modifies the contents of its random access memory (RAM); and a server's logs are updated and deleted at regular intervals when the law does not set a maximum duration for keeping them. In France, the length of time that ICT operators have to store connection data has been set at one year. Moreover, such evidence can be fragile owing to the storage medium: magnetic bands or optical devices sensitive to the environment, or error-prone read-write memory procedures. Even though such circumstances very seldom occur and are handled by error-correction procedures, the experts who interpret electronic data for courts of law have to be aware of them.

[1] This article has been translated from French by Noal Mellott (Omaha Beach, France).

In addition to new mediums and new environments, digital forensic investigators have had to adapt to many a change. They have had to learn how to collect evidence while systems are running — the problem of volatility has made it necessary to develop this so-called "live forensics". They have had to adjust to the growing uses of cryptographic techniques, whence the need to use deciphering methods. They have also had to learn to cope with the exponential growth in the volume of data being processed. They now have to collect evidence directly on computer networks.

In fact, the most important developments have involved collecting data on computer networks and the Internet by using the many new methods and tools necessary for having a handle on the various virtual "territories", *e.g.*, forums, social media, or even the secured or anonymous networks known as darknets (where investigators use anonymization techniques).

Furthermore, data are ever more often being exchanged with parties outside the inquest (with, for example, computer security firms or computer emergency response teams). This implies using the same formats for the data to be exchanged.

What ultimately matters is the ability to process, cross and analyze large volumes of data. Beyond collecting and analyzing data under conditions guaranteeing the integrity of the evidence, systems for processing big data are now gradually coming into use.

Changing standards and laws for digital evidence

Laws have gradually been adjusted to admit digital evidence in judicial inquests. The French penal code, for example, allows for copying data from an electronic medium during searches, conducting investigations using a pseudonym and remotely collecting data (THIÉRACHE & FREYSSINET 2018).

The next step, still under development at the European level, is to more efficiently exchange evidence among countries. For instance, it is currently possible to pool evidence for a joint investigation team within the European Union (with the support of Eurojust and Europol). Moreover, the European investigation warrant enables a judge to ask for investigatory operations (such as searches or requisitions) in others countries. For more efficiency in the simplest cases (*e.g.*, identifying an Internet user), an investigator in one country can make a binding request for information to an ICT operator in another country.

Standards are also an issue. Though not legally binding, ISO standard 17025:2005 on the requirements for testing and calibration laboratories is recognized at the European level, in particular by the European Network of Forensic Science Institutes (ENFSI). For application to digital forensics (so obviously different from the tests run in a laboratory of biology or chemistry), the differences in designing and validating the methods used have to be taken into account. In particular, it is important to be able to regularly update the software for computational forensics. The ICT laboratory of the forensic institute (IRCGN) of the French National Gendarmerie is accredited to run fourteen different tests.[2] There are other sources of legal standards, in particular for exchanging data with ICT operators. The European Telecommunications Standards Institute (ETSI) has drawn up about two dozen technical specifications on the conditions under which ICT operators are to make available to public authorities intercepted data and metadata.[3] These are the standards used by the national platform of lawful interceptions (PNIJ) in the French Ministry of Justice.

[2] The fourteen are listed on http://www.cofrac.fr/annexes/sect1/1-1916.pdf.
[3] See http://www.etsi.org/technologies-clusters/technologies/lawful-interception.

Criminal intelligence activities

As pointed out, data are now being collected and analyzed on a widespread scale during judicial inquests. Beyond digital and computational forensics as such, all key points of information in an investigation are being turned into data and processed, whether anthropometric information (of the sort collected by Bertillon in the 19th century), details of the inquest, or the results of the tests run on the evidence collected on the crime scene (FREYSSINET 2003). The processing of personal data related to a judicial inquest is, of course, subject to the applicable legislative and regulatory texts and, in France, to control by the National Commission on Informatics and Liberty (CNIL) or, in certain cases, a judge.

Traditional activities

Several types of data processing are used to establish relations between cases and inquests. Data are, for example, processed from criminal records (thus with the identification of suspects and a synopsis of convictions). For very important cases, the so-called serial databases (with more details) are used for more granular comparisons. In order to compare the evidence collected for a single, highly complex case, tools of criminalistic analysis serve to detect contradictions or simply make graphic descriptions of the relations between facts and persons.

Canadian law enforcement uses the services of academic research teams in criminology to obtain an outside view on their data: network analyses, the classification of behaviors or the validation of the methods used in an investigation.[4] In France, academic interventions are less current, but mention should be made of the MAPAP program for analyzing the circulation of pedophile images over peer-to-peer networks (FOURNIER & LATAPY 2015).[5]

Forensic activities

The information in the report filed by crime scene investigators and then from laboratory tests can also be entered into databases. Biometric databases already register fingerprints (physical or genetic), but several other types of information, less familiar to the general public, can also be entered in databases: ballistics (traces left by a gun on its bullets), evidence (such as ADN) on the tools used to open a door, or from ear lobes, shoe prints, insects, etc.

This stock of information is sometimes needed for a more granular description of the crime (*e.g.*, to locate the area from which dirt of a certain composition comes) or to relate facts (a drug and the person who supplied it). Using forensic databases of this sort calls for detailed analytics and for methods that make comparisons and relate facts to each other. Several developments are still possible.

Dematerializing penal procedures

The changes to come in judicial inquests will entail fully “dematerializing” penal procedures, from the complaint filed by a victim to the hearing before a judge, not to forget the documents produced by judges and investigators. Nowadays, these documents mainly circulate on paper (often with a full signature or initials on each page). On 10 January 2018, the ministers of Interior and of Justice jointly announced a major program for dematerializing penal procedures so that all parties concerned might have online access to all documents pertaining to their case.[6] This availability in electronic format of all the documents of an inquest opens new perspectives for both investigators and the lawyers representing plaintiffs and defendants.

[4] *Cf.* the studies by the team headed by Martin Bouchard at Simon Fraser University in Vancouver.
[5] http://antipaedo.lip6.fr/
[6] http://www.justice.gouv.fr/la-garde-des-sceaux-10016/dematerialisation-des-procedures-penales-31168.html

Exercising command over the data

Trends in technology, in particular big data and artificial intelligence, lead us to imagine going much farther in processing data for the purpose of fighting against criminality. In France, several programs of computer-assisted decision-making are in the works for comparing data from judicial investigations and from other sources (geographical, sociological, economic, meteorological…) (CREOGN 2017). This information could then be compared with police records (for example, the nature, place and date of police operations) in order to identify the most efficient measures to adopt, orient decision-making and analyze feedback objectively.

Several other developments have been imagined for improving the data processing of images or for comparing traces and evidence thanks to the use (and advances to be made) in artificial intelligence. The work of investigators and judges would be facilitated by tools for systematically exploring all the hypotheses, probable or less so, that they do not normally have the time to formulate. Such tools could also remind investigators of useful leads or other points of information.

The success of these programs does not depend on technical factors alone. This work can only be conducted thanks to efficient exchanges of data with all concerned parties (manufacturers, local authorities, researchers). It is necessary to make the technical, regulatory, ethical and, eventually, financial arrangements needed to develop these projects.

References

CREOGN (Centre de Recherche de l'École des Officiers de la Gendarmerie Nationale) (2017) "Hyperconnexion et résilience", *Revue de la gendarmerie*, 260, pp. 146-175.

FOURNIER R. & LATAPY M. (2015) "Temporal patterns of pedophile activity in a P2P network: First insights about user profiles from big data", *International Journal of Internet Science*, 10(1), pp. 8-19.

FREYSSINET É. (2003) "La preuve numérique: Un défi pour l'enquête criminelle du 21e siècle", *Les Cahiers du numérique*, 4(3), pp. 205-217.

THIÉRACHE C. & FREYSSINET É. (2018) "La procédure pénale face aux évolutions de la cybercriminalité et du traitement de la preuve numérique", report of the CECyF-Cyberlex work group of 24 January, 46p. Available via: https://www.cecyf.fr/wp-content/uploads/2018/01/Procédure-pénale-et-cybercriminalité-CECyF-Cyberlex.pdf.