Les données au cœur de la lutte contre la délinquance

Par Éric FREYSSINET
Colonel

Dès la fin du XIXe siècle, la collecte de données s'est imposée avec Alphonse Bertillon comme une des clés de la réussite des enquêtes judiciaires. Au XXIe siècle, la collecte, l'analyse et la présentation des données comme preuves au procès pénal sont au cœur de la lutte contre la délinquance. Elles se concrétisent dans le champ de la criminalistique numérique et du renseignement criminel et leur plein développement passera par une véritable maîtrise des données.

La criminalistique numérique

La criminalistique numérique (ou sciences forensiques numériques) regroupe l'ensemble des techniques de preuve utilisées dans les enquêtes judiciaires et reposant sur des données et supports numériques. Au passage, il est important de constater qu'il existe une forte proximité entre ces méthodes et celles utilisées par les personnes menant des investigations dans le domaine de la sécurité des systèmes d'information.

Évolutions technologiques de la criminalistique numérique

Émergents au début des années 1980, les éléments de preuve numérique sont désormais potentiellement omniprésents, quelle que soit la nature de l'infraction, avec le développement des techniques et des usages. Ils se présentent la plupart du temps sur des supports matériels, mais peuvent aussi être collectés dans l'environnement électromagnétique (comme pour la détection de communications Wi-Fi dans le domicile d'un suspect) ou encore être récupérés chez des tiers (parce qu'hébergés sur des serveurs ou détenus par un opérateur de communications électroniques).

Ces éléments de preuve numériques se caractérisent en outre par leur volatilité et leur potentielle fragilité. Ainsi, le contenu d'une mémoire vive informatique est constamment modifié par le système en fonctionnement, les journaux d'activité d'un serveur informatique font l'objet de rotations et sont régulièrement effacés, quand ce n'est pas la législation qui impose des durées maximales de conservation [1]. Enfin, ils peuvent être fragilisés par la nature des supports de stockage : supports magnétiques et optiques sensibles à leur environnement, ou encore mémoires informatiques susceptibles de générer des erreurs de lecture et d'écriture. Même si ces circonstances sont extrêmement rares et prises en compte par des dispositifs de correction d'erreur, elles doivent être connues de ceux qui interprètent les données numériques dans les enquêtes judiciaires.

Outre les nouveaux supports et les nouveaux environnements, les spécialistes de l'investigation numérique ont dû s'adapter à plusieurs évolutions : la nécessité de collecter des preuves sur des systèmes en fonctionnement (techniques dites de *live forensics*, nécessaires pour les questions de volatilité évoquées plus haut), le développement de l'usage des techniques cryptographiques (et la nécessité qui en découle de mettre en œuvre des méthodes de déchiffrement voire de cryptana-

(1) En France, la durée de conservation des données de connexion par les opérateurs de communication électronique est fixée à un an.

lyse), l'augmentation exponentielle des volumes de données à traiter et la collecte d'éléments de preuve directement sur les réseaux.

C'est bien ce domaine de la collecte de données sur les réseaux et sur Internet qui fait l'objet des développements les plus importants avec l'émergence de nouvelles méthodes et de nombreux outils qui s'avèrent nécessaires pour appréhender les différents territoires qui constituent Internet, tels que les forums et réseaux sociaux, ou encore les réseaux sécurisés ou anonymes autrement appelés *darknets* où les enquêteurs utilisent des techniques d'anonymisation.

Ensuite, l'échange de données avec des acteurs externes aux services d'investigation judiciaire, telles des entreprises de sécurité informatique ou des équipes de réaction aux incidents informatiques (CSIRT), devient de plus en plus courant, ce qui suppose d'utiliser des formats d'échange communs.

Au final, c'est la capacité à traiter, croiser et analyser des volumes de données importants qui devient la plus prégnante. Et donc au-delà de la collecte dans des conditions garantissant l'intégrité des éléments de preuve et l'analyse simple de ces données, ce sont de véritables systèmes de traitement de mégadonnées (*Big Data*) qui sont progressivement mis en œuvre.

Évolutions juridiques et normatives du traitement de la preuve numérique

La législation s'est progressivement adaptée pour accepter les éléments de preuve numériques dans les procédures judiciaires. Ainsi le Code de procédure pénale français (Thiérache et Freyssinet, 2018) prévoit la possibilité de copier des données depuis un support original au moment de la perquisition, de réaliser des enquêtes sous pseudonyme ou encore de collecter des données à distance.

L'étape suivante – en cours de développement au niveau européen – est de pouvoir échanger plus efficacement des éléments de preuve entre pays. Ainsi, il est aujourd'hui possible de mettre en commun des éléments de preuve dans le cadre d'une équipe commune d'enquête au sein de l'Union européenne (avec le soutien d'Eurojust et d'Europol). De même, l'ordre d'enquête européen permet à un magistrat de demander la réalisation d'opérations d'enquête (comme des perquisitions ou des réquisitions à des opérateurs) dans d'autres pays. Pour être encore plus efficace, dans les cas les plus simples (identification du titulaire d'un abonnement Internet par exemple), des demandes contraignantes pourraient être adressées d'un enquêteur dans un pays A à un opérateur de communications électroniques dans un pays B.

Les enjeux sont aussi normatifs. Ainsi, même si elle n'est pas légalement contraignante, la norme ISO 17025:2005 qui applique l'assurance qualité aux laboratoires d'essais est reconnue au plan européen (en particulier par l'association ENFSI des laboratoires de criminalistique) comme devant s'appliquer dans ces laboratoires. Pour s'appliquer au domaine numérique – manifestement différent des travaux réalisés dans un laboratoire de biochimie par exemple – il a fallu prendre en compte des spécificités quant à la construction et la validation des méthodes. En particulier, il est important de pouvoir régulièrement mettre à jour les logiciels utilisés par les experts judiciaires numériques. Le laboratoire informatique-électronique de l'Institut de Recherche criminelle de la Gendarmerie nationale (IRCGN) en France est accrédité [(2)] pour quatorze natures d'essais différents. D'autres normes trouvent à s'appliquer dans le champ de l'investigation numérique, notamment en matière d'échange d'informations avec les opérateurs. Ainsi, une vingtaine de spécifications techniques [(3)] produites par l'organisation européenne de normalisation ETSI précisent les conditions dans lesquelles le contenu des interceptions judiciaires et les métadonnées sont mis à disposition des autorités par les opérateurs. Ce sont ces standards qui sont mis en œuvre en France dans le cadre de la Plateforme nationale d'Interceptions judiciaires (PNIJ) gérée par le ministère de la Justice.

(2) http://www.cofrac.fr/annexes/sect1/1-1916.pdf
(3) http://www.etsi.org/technologies-clusters/technologies/lawful-interception

Le renseignement criminel

Comme nous l'évoquions en introduction, la collecte et l'analyse de données est réalisée plus globalement dans l'ensemble des enquêtes judiciaires. Ainsi, au-delà des éléments de preuve numérique, ce sont l'ensemble des éléments-clés de l'enquête qui peuvent être transformés en données et exploités, qu'il s'agisse d'informations anthropométriques telles que les collectait Bertillon au XIX^{e} siècle, de détails de l'enquête ou de données issues des analyses d'éléments collectés sur une scène de crime. Bien évidemment, s'agissant de données personnelles liées à une enquête judiciaire, ces traitements font l'objet de textes législatifs et réglementaires les encadrant et d'un contrôle de la CNIL, voire d'un magistrat référent dans certains cas.

Le renseignement criminel traditionnel

Plusieurs types de traitements de données sont mis en œuvre pour réaliser des rapprochements entre enquêtes judiciaires. En France, on trouvera par exemple le traitement des antécédents judiciaires (TAJ) avec les identités des personnes mises en cause et une synthèse des faits incriminés, et pour les phénomènes les plus importants des bases dites sérielles, contenant sur les affaires plus de détails permettant d'opérer des rapprochements plus fins. Au sein d'une même affaire particulièrement complexe, afin de confronter entre eux les indices collectés, d'identifier les contradictions ou simplement de réaliser des synthèses graphiques des relations entre les faits et les acteurs, des outils dits d'analyse criminelle sont mis en œuvre.

Les services de police canadiens font souvent appel à des équipes de recherche universitaire en criminologie pour apporter un regard extérieur sur leurs données : analyse de réseaux, classification des comportements ou encore validation des méthodes d'enquête [4]. En France, ces démarches académiques sont moins courantes, mais on pourra par exemple saluer les travaux du projet MAPAP [5] sur l'analyse de la diffusion des images pédopornographiques *via* les réseaux pair-à-pair (Fournier et Latapy, 2015).

Le renseignement forensique

Plus spécifiquement, les éléments issus des relevés réalisés par les techniciens sur la scène de crime, puis des examens de laboratoires criminalistiques, peuvent être intégrés dans des bases de données pour réaliser des rapprochements. C'est traditionnellement le cas des bases de données biométriques (empreintes digitales, empreintes génétiques), mais c'est aussi possible pour de nombreux autres types d'information moins connus dans le grand public : balistique (traces laissées par une arme sur un projectile), outils (empreinte laissée par les outils utilisés pour ouvrir une porte), lobes d'oreille (qui peuvent laisser une empreinte lorsqu'ils sont collés contre une surface), chaussures ou encore insectes, ADN de cannabis…

Ces bases de connaissances sont parfois nécessaires pour affiner les circonstances d'un fait criminel (par exemple retrouver le lieu géographique auquel correspond une certaine composition de terre) ou encore réaliser des rapprochements entre les faits (identifier le fournisseur d'un produit stupéfiant). La mise en œuvre de telles bases de données forensiques suppose à la fois de combiner des capacités analytiques suffisamment fines et de mettre en œuvre des méthodes de rapprochement efficaces : de nombreux développements sont encore possibles.

(4) On pourra notamment se rapporter aux travaux de l'équipe dirigée par Martin Bouchard de l'Université Simon Fraser de Vancouver.
(5) http://antipaedo.lip6.fr/

La dématérialisation de la procédure pénale

L'évolution suivante du renseignement judiciaire passera par la dématérialisation complète de la procédure pénale depuis la plainte de la victime jusqu'à l'audience devant le tribunal en passant par les pièces de procédure réalisées par les magistrats et les enquêteurs. Aujourd'hui, ces documents circulent essentiellement sous forme imprimée (et souvent signés ou paraphés à chaque page). Le 10 janvier 2018, les ministres de l'Intérieur et de la Justice ont conjointement annoncé [6] le lancement d'un grand projet de dématérialisation de la procédure pénale qui devra notamment permettre à tous les acteurs de la chaîne judiciaire d'accéder en ligne à un dossier unique. Cette mise à disposition de l'ensemble des pièces de l'enquête sous forme numérique ouvre de nouvelles perspectives, tant pour les acteurs de l'enquête judiciaire que pour les avocats des différentes parties.

Vers la maîtrise des données

L'évolution des technologies – notamment celles liées aux méga-données et à l'intelligence artificielle – nous permet d'envisager d'aller beaucoup plus loin dans l'exploitation des données pour la lutte contre la délinquance. Plusieurs projets d'analyse décisionnelle sont ainsi lancés en France (CREOGN, 2017), qui proposent de rapprocher les données issues des enquêtes judiciaires et les données provenant d'autres sources d'information complémentaires (géographiques, sociologiques, économiques, météorologiques, etc.). Elles pourraient ensuite être confrontées aux informations liées à l'action des services répressifs (par exemple la nature, le lieu et la date des opérations de contrôle réalisées) pour permettre d'identifier les mesures les plus efficaces, d'orienter la prise de décision et de favoriser une analyse objective du retour d'expérience.

De nombreux autres axes de travail sont envisagés, pour améliorer le traitement des images et le rapprochement des traces et indices, grâce à l'intégration ou l'amélioration de l'intelligence artificielle. On peut aussi imaginer de faciliter le travail des enquêteurs et des magistrats pour explorer de façon plus systématique l'ensemble des hypothèses probables ou moins probables qu'ils n'auraient pas le temps d'envisager dans un temps raisonnable et leur permettre de n'oublier aucun élément ou piste utile.

La réussite de tels projets ne repose pas que sur des facteurs techniques. Ce travail ne pourra être réalisé que grâce à des échanges de données efficaces avec tous les acteurs concernés (industriels, collectivités locales, chercheurs) : il faut inventer les cadres techniques, réglementaires, éthiques et éventuellement financiers nécessaires à leur mise en place.

Bibliographie

CREOGN – Centre de recherche de l'école des officiers de la Gendarmerie nationale (2017), « Hyperconnexion et résilience », *Revue de la Gendarmerie nationale*, n° 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 XXIe 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 », CECyF & Cyberlex.

(6) http://www.justice.gouv.fr/la-garde-des-sceaux-10016/dematerialisation-des-procedures-penales-31168.html