

INTERNATIONAL >
Cybersécurité et régulation
internationale

DÉFENSE >
Résilience des SIC militaires

DROIT >
Du cybercrime au cyberjuge

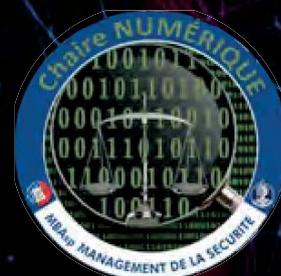


REVUE

de la gendarmerie nationale

REVUE TRIMESTRIELLE / DÉCEMBRE 2017 / N° 260 / PRIX 6 EUROS

Hyperconnexion
et résilience



DOSSIER

Hyperconnexion et résilience

Nouvelles complexités, nouvelles menaces 53 par Gilles Hilary	Analyse économique des monnaies virtuelles 97 par Jean-Luc Delangle
Une convention de Genève pour le numérique ? Non ! 57 par Anne-Thida Norodom	CyberEdu, parler de sécurité numérique dans les cours 115 par Gérard Peliks
WannaCry et la diffusion des zero day exploits 61 par Gilles Hilary	Les enjeux de l'hyperconnexion : de la smart à la safe cities 121 par Myriam Quéméner
La notation de cybersécurité 67 par Guillaume Tissier	État des lieux de la sécurité des objets connectés 127 par Cyril Nalpas
Mieux vaut guérir que prévenir 75 par Didier Danet	Cybermalveillance.gouv.fr C'est parti ! 133 par Jérôme Notin
Data Stratégie 81 par François Cazals	La donnée, nouvelle préoccupation du comité exécutif 139 par Gérard Hatabian
La cybersécurité "marétique", bilan et perspectives 87 par Michel Bénédictini	
Cybersécurité maritime : le cap est donné ! 91 par Barnabé Watin-Augouard	

Data

Stratégie

Par **FRANÇOIS CAZALS**

L

La démocratisation totale de l'Internet, du téléphone mobile et des machines connectées fait du traitement des données un enjeu stratégique pour les organisations, en général, et la gendarmerie nationale, en particulier. Nous analyserons ici ce véritable phénomène des données omniprésentes et comment il est possible de les valoriser pour augmenter la qualité des stratégies et des organisations. Nous envisagerons finalement quelques pistes de réflexion qui s'appliqueraient spécifiquement pour la gendarmerie nationale.



FRANÇOIS CAZALS

Professeur adjoint à HEC Paris - Consultant en stratégie - Lieutenant-colonel (réserve citoyenne) cabinet DGGN

Un monde de données

L'explosion de la production de données dans le monde est un phénomène relativement récent qui commence réellement en 2012. Avec une dé-

mocratisation réelle de l'Internet (50 % de personnes connectées au niveau mondial), des médias sociaux (2 milliards d'inscrits actifs sur Facebook), du téléphone mobile (autant de cartes SIM que d'humains sur la planète) et la montée en puissance des objets connectés (50 milliards en 2017), ce phénomène va s'inscrire dans la durée. Cette génération massive et exponentielle de données fonde le phénomène *Big data*.

Évidemment, la possibilité pour chacun d'accéder très facilement à toutes ces données transforme profondément la nature même de nos comportements et crée de nouveaux modèles économiques. Google a déjà théorisé le fait que la consultation d'Internet s'inscrit désormais en amont de tous nos processus de décision, professionnels ou personnels, au travers de son modèle ZMOT (*Zero Moment of Truth*)¹. Parallèlement, nous voyons également émerger une véritable économie des plateformes numériques,

(1) Le ZMOT évoque donc le fait que le consommateur utilise désormais les moteurs de recherche, les réseaux sociaux et les mobiles pour obtenir des avis ou informations sur le produit/service pour lequel il a reçu un stimulus. La personne initie cette pratique au moment de la recherche d'informations mais également après que sa décision d'achat ait été prise pour donner à son tour un avis suite à l'utilisation du produit/service.

qui disruptent avec violence les modèles existants. Ainsi, les Uber, AirBnB et autres Netflix modifient profondément les équilibres économiques et sociaux de pans entiers des industries traditionnelles. Cette nouvelle étape du développement de l'économie numérique se traduit de manière spectaculaire dans les valorisations boursières des entreprises internationales. Ainsi, six entreprises

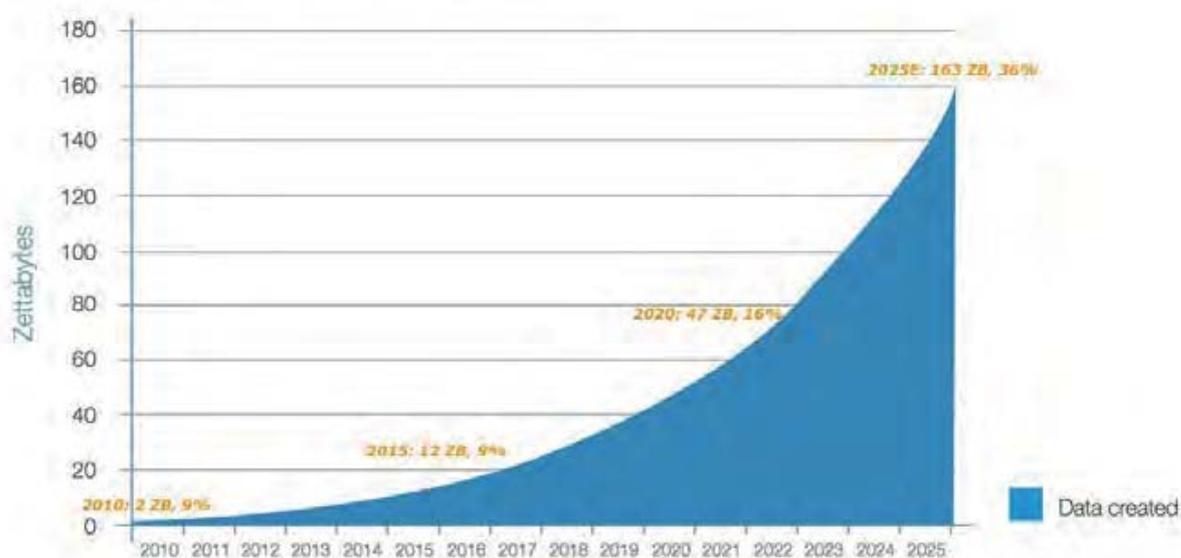
du numérique figurent parmi les dix premières capitalisations planétaires. Cette numérisation du monde représente une

véritable révolution dont nous ne vivons que les prémices et la convergence des technologies NBIC (nanotechnologies, biotechnologies, technologies de l'information et technologies cognitives) ne va que l'amplifier. Il va de soi que cette nouvelle donne est porteuse autant d'espoirs que de dangers. La croissance des cyberattaques, la criminalité numérique, l'utilisation des technologies par les organisations terroristes et les réseaux cryptés (*DarkNets*, messageries cryptées) posent évidemment de nombreux défis très préoccupants pour la sécurité publique.

Créer de la valeur avec les données

L'enjeu principal pour les organisations est de tirer profit de ces gisements de données et de les valoriser. Mais de quelles

Figure 2. Annual Size of the Global Datasphere



Source: IDC's Data Age 2025 study, sponsored by Seagate, April 2017

source idc via @mikequindazzi

données parle-t-on ? En effet, celles-ci sont très loin d'être homogènes. Si l'on reprend la métaphore bien connue des données constituant le « *pétrole du XXI^e siècle* », celle-ci suggère avec justesse que les données brutes ne sont pas « raffinées » et exigent d'être purifiées pour pouvoir être exploitées. De manière plus précise, on distingue trois grandes catégories de données. Les données structurées, ou données opérationnelles, sont générées par les organisations et stockées dans leurs systèmes d'information. Il s'agit essentiellement de données de gestion. Elles ne représentent qu'environ 10% du phénomène *Big data*. La majorité des données sont non-structurées ou semi-structurées (étiquetées par une ou plusieurs variables descriptives). Il s'agit de textes, de sons, d'images et de vidéos. Ces données sont issues d'Internet (sites Web et messageries), des médias sociaux et des machines connectées.

Valoriser ces données si variées consiste à relever un quintuple défi. Le volume des données, en expansion exponentielle, pose un défi en termes de captation et de stockage. La vitesse de production et de circulation des données amène à devoir également traiter et analyser les données très rapidement. La variété des données (structurées ou brutes, chiffres, textes, images, son, vidéos, ...) rend les traitements

et analyses nettement plus complexes.

La valeur des données est très difficile à apprécier, au plan de leur qualité intrinsèque et de leur véracité. La visualisation constitue le dernier enjeu critique en termes d'aide à la décision : comment passer de traitements analytiques complexes à des restitutions suffisamment simples et explicites pour les décideurs ?

L'évolution des technologies constitue donc un point clé du développement des Data Stratégies. Aujourd'hui, il est possible et assez économique de capter, stocker et traiter de très grandes masses de données grâce à de nouvelles infrastructures informatiques dites distribuées. Les plus connues des plateformes, disponibles en Open Source, sont Hadoop (stockage distribué) et Apache Spark (traitements distribués) qui peuvent être associées harmonieusement. La puissance des matériels a parallèlement évolué de manière spectaculaire, avec des solutions relativement accessibles au plan financier. À titre d'illustration, le nouveau super calculateur NVIDIA DGX-1, capable

RANK	APRIL 2017	Q4 2011	Q4 2006
1	Apple: 741	Exxon Mobil: 406	Exxon Mobil: 447
2	Alphabet: 585	Apple: 376	General Electric: 384
3	Microsoft: 505	PetroChina: 277	Microsoft: 294
4	Amazon: 432	Royal Dutch Shell: 237	Citigroup: 274
5	Facebook: 408	ICBC: 228	Gazprom: 271
6	Berkshire Hathaway: 404	Microsoft: 218	ICBC: 255
7	Exxon Mobil: 344	IBM: 217	Toyota: 241
8	Johnson & Johnson: 330	Chevron: 212	Bank of America: 240
9	JPMorgan Chase: 303	Walmart: 205	Royal Dutch Shell: 226
10	Alibaba Group: 278	China Mobile: 196	BP: 219

Legend: Data-driven company

Source: S&P Capital IQ, "Top 10 Companies with Highest Market Capitalization Worldwide."
Note: Market capitalization figures have been rounded and are in \$billions.

de fournir une puissance de traitement de 170 Tflops (170 000 milliards d'opérations par secondes), coûte moins de 130 000 \$.

Concernant les traitements des données proprement dits, un socle très robuste de modèles informatiques, analytiques et mathématiques existe déjà. L'informatique décisionnelle (*Business Intelligence*) permet depuis longtemps d'analyser les données opérationnelles des organisations. De nombreux modèles statistiques explicatifs et prédictifs (*Data Mining*) permettent d'en extraire une véritable valeur opérationnelle.

Les technologies d'apprentissage automatique des ordinateurs (*Machine Learning*) augmentent la puissance et la complexité des traitements. Si elles sont déjà anciennes, un véritable bond conceptuel et technologique a été réalisé dans les années 2010 avec les modèles d'apprentissage profond (*Deep Learning*). Le chercheur français Yann LE CUN, aujourd'hui directeur de l'intelligence artificielle chez Facebook, en a été un des pionniers principaux au niveau mondial. Grâce aux réseaux de neurones artificiels profonds ainsi imaginés, la résolution de traitements très complexes est possible. L'ensemble de ces progrès technologiques permet le développement d'applications d'intelligence artificielle, qui constituera certainement un des piliers principaux des Data Stratégies, à l'avenir.

Concevoir et déployer une Data Stratégie

Cela semble trivial, mais élaborer une Data Stratégie découle évidemment du mana-

gement stratégique d'une organisation et d'une vision de direction générale. Or, la tentation est grande de considérer que la valorisation des données ne relève que de considérations technologiques, en général, et surtout informatiques, de manière plus spécifique. Ainsi, il est fréquent de rattacher des entités de Data Stratégie (souvent nommées *DataLab*) à la direction des systèmes d'information. Le risque est grand, dans cette configuration, de multiplier les tests (*Proof of Concept* ou POC), sans qu'une ligne directrice n'existe et avec des taux d'échecs importants.

Les organisations les plus avancées adoptent une démarche très différente. Il s'agit, au préalable, de cartographier les données dans les différentes structures managériales, de définir les principes et règles de base d'accès et de traitement et ce en conformité avec les lois et règlements existants. Ces fonctions centrales de gouvernance des données et des traitements sont généralement confiées à une fonction de direction de l'entreprise : la direction du pilotage stratégique des données. Les Anglo-Saxons ont imaginé une dénomination spécifique pour le dirigeant d'une telle structure : *le Chief Data Officer*.

Au-delà des règles et principes liés aux données, la direction générale de l'organisation doit définir un certain nombre d'axes stratégiques applicables aux domaines où les données semblent susceptibles de dégager des gisements de valeur : meilleure connaissance opérationnelle, potentiels d'économie et d'efficacité managériale, alertes critiques,

pilotage plus fin et en temps réel des performances... Ces orientations doivent procéder d'une démarche participative avec les grandes directions métiers de l'organisation, qui doivent en constituer les parties prenantes principales au niveau opérationnel. Il est notable de constater que les créations de valeur concernent potentiellement toutes les fonctions régaliennes d'une organisation : direction générale, évidemment, mais également direction du marketing et de la relation client (ou usagers, voire citoyens, dans des contextes publics), direction financière, directions opérationnelles (production, logistique, distribution...) et direction des ressources humaines.

Sur le fondement d'orientations stratégiques claires, au service des métiers de l'organisation, une démarche agile doit être imaginée pour dégager les potentiels de valeur cibles. L'observation des meilleures pratiques actuelles suggère une démarche en trois macro-étapes.

La première phase est celle de l'expérimentation. Elle consiste à développer un prototype de test, pour résoudre une problématique donnée. Une équipe pluridisciplinaire est nécessaire, à ce stade, pour évaluer l'environnement dans sa globalité : experts techniques des données (*Data Engineers*), spécialistes de la modélisation informatique et statistique (*Data Scientists*) et spécialistes métiers. Lorsque le modèle *In Vitro* dans la structure d'étude (DataLab) est valide, il convient de le confronter à la réalité *In Vivo* pour vérifier sa robustesse opérationnelle.

La dernière étape n'est pas la moindre puisqu'il s'agit d'intégrer le modèle dans une logique de déploiement industriel. Cette étape cruciale nécessite évidemment une adaptation de certains processus de l'organisation et une véritable intégration au système d'information.

Quelle Data Stratégie pour la gendarmerie nationale ?

L'Institution s'est naturellement engagée dans la voie de la valorisation de ses données, compte tenu des nombreux enjeux qui en découlent.

Le premier enjeu est opérationnel et consiste à pouvoir accéder aux données « Métiers », en temps réel, dans le contexte de mobilité consubstantiel aux missions des gendarmes (prévention, sécurité publique « à chaud »). Il concerne également les unités de commandement (DGGN, régions, groupements, ...) pour affiner le pilotage de la performance, et être alerté, en temps réel ou quasi-réel, des informations critiques issues du terrain.

Le second enjeu est économique. Il s'agit d'identifier des gisements de valeur financière et d'optimiser l'efficacité opérationnelle, par un management prédictif, la numérisation et l'automatisation de certains processus.

Le troisième enjeu est relationnel. Il intéresse, en temps réel, la surveillance de l'e-réputation de la gendarmerie nationale, notamment en période de crise, et l'orientation de sa communication institutionnelle

grâce, en particulier, à l'analyse des données sémantiques du Web et des médias sociaux.

Ces enjeux ont évidemment été perçus par la DGGN et plusieurs initiatives matérialisent la structuration de la Data Stratégie de l'Arme. Une structure dédiée à l'expérimentation des projets de traitement de données existe déjà : le DataLab. Elle est rattachée à la direction commune des systèmes d'information de la gendarmerie nationale et de la police nationale (ST[SI]2). Plusieurs projets pilotes ont été menés concernant à la fois des sujets liés à la sécurité publique (modèles prédictifs liés à la délinquance) et à l'efficacité opérationnelle (maintenance préventive de la flotte automobile, par exemple).

Au niveau de la DGGN, une entité spécifique définit la stratégie numérique de l'institution : la mission numérique. Son impact sur la transformation du modèle opérationnel de la gendarmerie nationale est évident, notamment au travers du déploiement en mobilité des applications Métiers avec NEOGEND, qui constitue « le nouvel assistant du gendarme » ! Une gouvernance stricte des données et des algorithmes est évidemment nécessaire, compte tenu des évolutions juridiques et surtout de l'application dès le mois de mai 2018 du règlement général pour la protection des données. Ces sujets font déjà l'objet de réflexions et d'études, comme le démontre le dernier atelier du centre de recherche de l'EOGN, le 26 septembre 2017 sur le thème des « Algorithmes prédictifs : quels en sont les enjeux éthiques et juridiques ? ».

Ces premières initiatives démontrent l'importance du développement d'une Data Stratégie structurée pour la gendarmerie nationale. Il ne s'agit pourtant que des premières étapes d'un plan plus ambitieux.

Intercepter des signaux faibles de l'environnement, réagir avec agilité et en temps réel à l'imprévu, optimiser les moyens et ressources, valoriser l'image institutionnelle, restituer au citoyen une information utile : voici quelques axes de développement évidents pour renforcer l'efficacité de l'institution et la relation avec les citoyens, à partir d'une stratégie de valorisation de ses données.

Au-delà de la vision stratégique et des contingences technologiques, une véritable transformation de l'organisation et un renforcement des compétences doivent être imaginés. La formation initiale et continue des gendarmes sur les thèmes du numérique et des données constituera certainement un facteur clé de succès de la conduite du changement nécessaire.

L'AUTEUR

François CAZALS est professeur adjoint à HEC Paris. Il y enseigne la stratégie et le marketing et s'est spécialisé sur les sujets des stratégies et de la transformation numériques et les stratégies de valorisation des données grâce à l'intelligence artificielle. Dirigeant d'un cabinet de conseil sur ces thèmes, il est également auteur de nombreux livres et articles, notamment « Stratégies digitales » (De Boeck, 2015). Il intervient régulièrement pour la DGGN et le centre de recherche de l'EOGN, en qualité de lieutenant-colonel (réserve citoyenne) affecté au cabinet du directeur général