L'enjeu des données pour la cyberdéfense

Par Didier DANET
Maître de conférences (HDR) de l'Université de Rennes 1

L'espace numérique est un champ de conflictualité, où la maîtrise des données est un enjeu dont l'importance ne fait que croître. En effet, contrairement à une vision trop répandue, la cyberdéfense ne s'intéresse pas uniquement à la protection des systèmes d'information interconnectés (le contenant), mais s'intéresse tout autant à celle des contenus informationnels, dont les données forment la matière première. Mais si la maîtrise des données devient ainsi un enjeu central de la cyberdéfense, elle est rendue presque illusoire, notamment du fait de la « datafication » du monde, de la révolution numérique à l'œuvre et de nouvelles pratiques sociales affaiblissant la capacité de contrôle des États sur la production et la circulation des données. Dans cet article, nous suggérons d'approfondir deux pistes de réflexion : la première vise à mieux protéger les stocks de données qui sont en possession des institutions civiles et militaires, et la seconde à responsabiliser les acteurs dans la génération des flux de données.

L'espace numérique est un champ de tensions et de conflits. Les rivalités économiques, politiques, culturelles… s'y déploient, comme elles le font dans les autres champs de la vie sociale. Pour les militaires, l'espace numérique est un champ immatériel, dans lequel les armées doivent être en mesure de se défendre et de s'affirmer (Douzet et Géry, 2020). Tel est l'objet de la cyberdéfense, devenue un domaine prioritaire au regard des investissements et des recrutements opérés par le ministère des Armées depuis plusieurs années (Taillat *et al*., 2018).

Mais, dans ce cadre général de la conflictualité numérique, quel est l'enjeu des données pour la cyberdéfense ?

Les données, un enjeu pour la cyberdéfense

Il serait excessif de dire que la maîtrise des données a été le point de départ du processus qui a abouti à faire de la cyberdéfense un élément prioritaire de la Défense française. La préoccupation initiale était plus la protection des systèmes d'information (le contenant) que de l'information ou des données elles-mêmes (le contenu). Mais il est vite apparu qu'il était impossible de traiter du contenant sans s'inquiéter du contenu, ce qui a conduit à faire des données un enjeu essentiel de la cyberdéfense.

Les données et les différentes couches de l'espace numérique

Parmi les nombreuses missions dont est chargée la cyberdéfense, certaines relèvent de la protection des systèmes d'information, dont l'interconnexion généralisée fait naître la crainte d'une attaque aux effets dévastateurs, le Pearl Harbor numérique. L'enjeu est alors principalement technique : comment empêcher un assaillant de se servir d'une brèche dans l'architecture ou les systèmes d'exploitation du réseau pour y pénétrer et le paralyser ou le saboter ? Dans ce cas de figure, les données contenues dans le système à défendre forment un objet secondaire. L'attaquant a pour objectif de prendre le contrôle d'un système d'information pour s'en servir à son profit (attaque subie par TV5 Monde) ou en faire un outil d'espionnage qui se retourne ainsi contre ses utilisateurs (le virus informatique Flame visant, notamment, certains responsables iraniens). L'attaquant peut encore vouloir le paralyser (attaques lancées en 2007 contre l'Estonie) ou le détruire (Stuxnet employé contre les centrifugeuses iraniennes ; ou KillDisk contre les centrales électriques ukrainiennes à la fin 2015). Du point de vue de la cyberdéfense, la question centrale est bien de protéger l'intégrité du système d'information et de maintenir la connexion entre elles des différentes composantes du système : le défenseur doit empêcher l'intrusion ou, si elle se produit, la détecter, la cantonner et l'éradiquer. Cette dimension conserve toute son importance et son actualité est constamment renouvelée. C'est ainsi que l'invasion de l'Ukraine par la Russie a été précédée d'une attaque visant à détruire les modems utilisés par l'armée ukrainienne pour traiter les signaux passant par le satellite KA-SAT de la société américaine Viasat.

Mais si elles occupent une place relativement secondaire dans ce type d'attaques, les données forment au contraire l'objet principal d'actions qui visent à leur captation, à leur indisponibilité ou à leur destruction. Pour les actions d'espionnage qui se déploient dans l'espace numérique, le but est de s'introduire dans un système d'information pour y dérober des données qui seront ensuite exploitées à des fins diverses : connaître les projets d'un compétiteur, diffuser les éléments qui

le placeront en situation délicate (publication de *mails* confidentiels comme cela a été le cas à la suite du piratage de Sony en 2014 ou du Parti démocrate en 2016) ou encore rendre indisponibles en les chiffrant les données dont le détenteur a besoin pour conduire son activité (Crypto Locker, Petya, Wannacry). Les attaques cyber contre les données peuvent également avoir pour but la destruction pure et simple de ces données. Dans le cadre du conflit ukrainien, plusieurs attaques par "wipers" ont été recensées (HermeticWiper, IsaasWiter, CaddyWiper), la raison d'être de ces logiciels malveillants étant de provoquer l'effacement irrémédiable des données figurant sur les disques qu'ils parviennent à atteindre.

Enfin, les données sont la matière première principale des manœuvres informationnelles qui relèvent de la cyberdéfense. L'accès aux données de terrain est essentiel pour contredire les narratifs adverses ou affirmer les siens. Il a été ainsi relevé que les autorités américaines avaient innové en diffusant largement les données obtenues par les services de renseignement quant à l'accumulation de moyens militaires russes aux frontières de l'Ukraine, établissant ainsi la volonté préméditée de la Russie d'envahir un État souverain aux frontières internationalement reconnues et faisant d'elle, par là même, l'agresseur aux yeux de l'opinion internationale. De même, l'exploitation d'images satellites produites par des compagnies privées et librement accessibles a permis de lever le doute sur la responsabilité de l'armée russe dans les massacres de Boutcha. Il en va de même pour le recensement de fosses communes susceptibles de révéler des crimes de guerre.

Il apparaît donc nettement que la maîtrise des données est, au même titre que la protection des systèmes d'information interconnectés, une préoccupation essentielle de la cyberdéfense.

La maîtrise des données mise au défi de la massification des techniques de communication et des pratiques sociales

La multiplication à l'infini des objets qui captent ou génèrent des données et permettent leur circulation bouleverse les modes de gouvernement des sociétés avec l'émergence de « nouvelles formes d'expression du pouvoir qui se développent par le biais des outils numériques », ce que l'on a pu qualifier fort justement de processus de « datafication généralisée » (Cattaruzza, 2019). Parmi ces outils, les objets communicants, qui prolifèrent dans le domaine civil comme militaire, prennent une place de plus en plus importante et conduisent à la massification de la production de données de tous types. Cette massification soulève en premier lieu des questions intrinsèques de sécurité, c'est-à-dire que le prix modique de nombre de ces objets communicants ne permet pas de couvrir le coût de conception et de mise en place de véritables mesures de sécurité, ce qui laisse ces objets à la merci du premier cyber criminel venu. Surtout, cette massification interroge quant à la maîtrise des données (Danet et Desforges, 2021). En effet, si certaines d'entre elles peuvent ne pas être sensibles en apparence (contenu d'un réfrigérateur, kilométrage et heure d'utilisation d'une trottinette, etc.), d'autres le sont à l'évidence beaucoup plus (données de santé, prestations sociales, sécurité du domicile, par exemple). Ces données sont souvent générées sans même que nous n'en ayons conscience. Or, elles peuvent dévoiler des caractéristiques, voire des failles individuelles ou collectives, offrant ainsi d'énormes opportunités de ciblage à des fins commerciales ou d'espionnage visant des personnes ou des organisations civiles ou militaires.

La « datafication généralisée » doit beaucoup aux avancées de la technologie qui en sont la condition nécessaire. Mais elle n'en doit pas moins aux comportements individuels et aux pratiques sociales qui en forment le véritable moteur. Rien n'oblige véritablement le propriétaire d'une maison à l'équiper d'une sonnette connectée qui va envoyer vingt-quatre heures sur vingt-quatre et sept jours sur sept les images prises par cet équipement connecté sur un *cloud,* où l'opérateur du service, qui s'est réservé le droit d'un accès illimité à ces images, pourra les exploiter à son gré avec tous les moyens de l'intelligence artificielle dont il est l'un des spécialistes mondiaux.

Les conséquences de la « datafication généralisée » ne sauraient mieux être illustrées dans le monde militaire par le cas Strava. Au début de l'année 2018, un étudiant australien, Nathan Ruser, révèle qu'il a découvert l'existence de bases militaires secrètes en Afghanistan, en Syrie et au Niger. Il n'a eu besoin pour ce faire que d'exploiter les données librement disponibles sur Strava, une application permettant d'enregistrer des activités sportives à partir de montres connectées. En croisant le profil des usagers de l'application et les parcours enregistrés par ces derniers, il n'était guère difficile de donner du sens à la concentration de militaires faisant du sport dans des régions en guerre ou dans des endroits où il n'y aurait apparemment aucune infrastructure répertoriée.

Pour les forces armées comme pour les administrations civiles, un véritable changement de paradigme s'est donc produit avec la révolution numérique. La production et la circulation des données échappent désormais très largement à toute forme de contrôle ou de coercition, et la tendance ira en s'accroissant sous l'impulsion de la massification des objets communicants ainsi que de la participation de plus en plus large des individus à toutes les formes de communautés numériques. La masse des données exploitables, sensibles ou non, a donc vocation à se multiplier et à offrir à ceux qui sauront exploiter ces données la possibilité de prendre l'ascendant sur des individus précisément ciblés ou des groupes plus ou moins nombreux, voire des collectivités nationales entières.

Intégrer l'enjeu de la maîtrise des données dans les opérations militaires

L'émergence de l'espace numérique comme champ de conflits a conduit à concevoir et à mettre en place des actions spécifiques, complémentaires des opérations militaires plus conventionnelles. La guerre résultant

de l'invasion par la Russie de l'Ukraine joue comme un révélateur des enjeux actuels de la maîtrise des données dans une situation d'affrontement militaire ; il laisse entrevoir ce que ces enjeux pourraient devenir dans un avenir proche, rendant plus nécessaire que jamais une stratégie de formation à l'usage des outils numériques.

De l'Ukraine à la « guerre cognitive »

Les premiers enseignements de la guerre d'Ukraine ont montré toute l'importance de la maîtrise des données pour des belligérants qui cherchent tous à l'emporter dans le champ informationnel. Mais cette maîtrise pourrait revêtir un caractère encore plus décisif avec le développement du concept chinois de « guerre cognitive » ("intelligentized warfare"). De manière synthétique, ce concept renvoie à l'idée d'agir sur le processus de décision d'individus précisément ciblés (une autorité politique, des groupes appartenant à des forces spéciales…) ou, au contraire, de groupes plus larges de la population, afin de les amener à agir dans le sens des intérêts de leur adversaire (Takagi, 2022). Il s'agit donc moins d'une révolution dans la conduite des affaires militaires que du prolongement et de l'amplification des manœuvres d'influence traditionnelles. Le point intéressant concernant les données est que la clé technologique de la guerre cognitive est l'intelligence artificielle et que celle-ci ne saurait être efficace si elle ne disposait pas de très grands volumes de données sur les cibles visées, notamment des données personnelles de toute nature. Il en résulte une inquiétude au regard de certaines attaques cyber qui ont permis à leurs auteurs de s'emparer de très gros fichiers contenant ce type de données : la double attaque visant l'Office of Personal Management américain à l'été 2014, puis durant l'hiver 2015 ; le piratage des fichiers clients des chaînes d'hôtels Marriott et Hilton ou de ceux de la Marine américaine en 2016.

Les mesures à prendre

La maîtrise des données est donc un enjeu central pour la cyberdéfense. Quelles mesures celle-ci peut-elle s'efforcer de mettre en œuvre pour y répondre ? Deux axes d'effort nous semblent possibles : d'une part, protéger les stocks de données, notamment des données sensibles, que l'institution militaire doit pouvoir mobiliser pour remplir ses missions, et, d'autre part, endiguer les processus indésirables de production des données lorsque la cyberdéfense est en mesure de le faire.

Le renforcement de la protection des stocks de données

Toute institution, civile comme militaire, produit des données qu'elle structure afin de répondre à ses besoins. De cette banale réalité découle pour l'institution concernée une responsabilité particulière en ce qui concerne la protection de ses données contre toute forme d'altération, de captation ou de destruction, ce qui suppose *a minima* le contrôle de l'accès auxdites données. N'importe qui ne doit pas pouvoir accéder à n'importe quelle donnée stockée par une administration ou une entreprise, que cette donnée soit sensible ou non. Quatre questions relatives à cette protection nous semblent utiles à prendre en considération, non seulement pour la cyberdéfense mais aussi pour l'ensemble des acteurs concernés :

- La légitimité de l'ouverture des données publiques : le principe de l'*open data* offre aux acteurs les plus puissants, très souvent étrangers, un accès sans restriction à des données qui, pour être publiques, n'en sont pas moins susceptibles de fournir un atout dans la compétition à ceux qui disposent déjà d'avantages évidents en termes technologiques, commerciaux, industriels…
- La prévention des fuites accidentelles : la protection des données par les institutions qui en sont responsables suppose des mesures adaptées et proportionnées qu'il ne serait pas pertinent de laisser à leur seule libre appréciation. à l'image des règles instaurées en matière de données de santé ; des règles *ad hoc* pourraient être mises en place dans d'autres domaines ou pour d'autres acteurs considérés jusqu'ici comme moins sensibles.
- La lutte contre les pratiques prédatrices : le développement rapide des techniques et des outils du *Big Data* pousse certains prestataires à se spécialiser dans la collecte, la structuration et la revente des données. Certaines pratiques jusqu'ici légales interrogent sur leur légitimité ; une interdiction ou, à tout le moins, un contrôle de celles-ci pourraient être envisagés. C'est ainsi que des entreprises comme Lexis Nexis commercialisent des fichiers de données portant sur des militaires américains (Sherman, 2021). La menace représentée par ce type de produits apparaît assez évidente.
- La prévention de la dépendance à l'égard de prestataires étrangers : le domaine de la collecte et du traitement des données est largement dominé par les acteurs américains du numérique : Amazon ou Microsoft pour le stockage dans le *Cloud*, Palantir pour l'analyse… Certes, ces acteurs sont les plus avancés techniquement et sont les seuls en mesure d'offrir des services de qualité à des coûts très intéressants. Mais sans les accuser de n'être que le faux nez de l'administration américaine, leur positionnement stratégique et la relation qui peut être établie avec eux obligent à tenir compte du risque que leur nationalité représente.

La maîtrise de la génération des flux de données

La production et la circulation des flux de données semblent impossibles à endiguer tant les individus disposent aujourd'hui d'un large éventail de moyens techniques pour ce faire et tant la norme sociale les pousse à s'y adonner. L'empreinte numérique de chacun de nous s'étend constamment et, surtout, elle ne s'efface pas. S'agissant de l'institution militaire, il ne saurait être question pour elle de limiter ce risque en interdisant purement et simplement à ses personnels de participer à la vie numérique de la Nation. Il y aurait là une atteinte inadmissible aux libertés individuelles de plusieurs centaines de milliers d'agents publics. En outre, cette interdiction serait inopérante si elle ne s'appliquait pas également aux conjoints, aux enfants et, *in fine*, aux familles et proches de ces militaires, ce qui reviendrait à vouloir interdire tout accès aux réseaux sociaux à plusieurs millions de citoyens. Il faut donc

compter sur la responsabilité personnelle et collective des personnels militaires et de leur famille pour qu'ils adoptent une conduite responsable dans l'univers numérique. Un *post* ou un *tweet* pouvant mettre en péril une mission, voire la vie des hommes qui la mènent, il ne serait pas exagéré de former les militaires et leurs familles à un usage raisonné des outils numériques.

Conclusion

Après s'être longtemps concentrée sur les risques associés au sabotage des infrastructures et au piratage des logiciels ou des applications, la cyberdéfense a pris toute la mesure de l'enjeu que représente la maîtrise des données pour la conception et la conduite des opérations militaires. La production, la collecte, le stockage, l'exploitation et la diffusion des données forment dès aujourd'hui les facteurs clés du succès des organisations militaires fonctionnant en réseau. Les données seront demain au cœur du combat collaboratif ou des opérations d'influence, voire de la « guerre cognitive ». La guerre en Ukraine a montré à la fois l'importance cruciale de l'enjeu que représentent les données mais aussi la difficulté à les maîtriser, notamment au regard de la multiplicité des parties prenantes et du contrôle limité que les autorités étatiques peuvent exercer sur la production et la circulation des données.

Une politique de maîtrise des données est donc indispensable pour renforcer la résilience de la Nation face à la menace de conflits armés. Elle est en même temps très complexe à définir tant les leviers habituels de la puissance publique ont peu de prise sur l'objet qu'il s'agit de réguler.

Bibliographie

CATTARUZZA A. (2019), *Géopolitique des données numériques*, Paris, Le Cavalier Bleu.

DANET D. & DESFORGES A. (2021), « Des objets connectés aux objets communicants. Les enjeux de souveraineté des objets communicants », *Enjeux numériques – Annales des Mines*, n°16, décembre, pp. 81-85.

DOUZET F. & GERY A. (2020), « Le cyberespace, ça sert, d'abord, à faire la guerre. Prolifération, sécurité et stabilité du cyberespace », *Hérodote*, n°177-178, pp. 329-349.

SHERMAN J. (2021), "Data Brokers Are Advertising Data on US Military Personal", *Lawfare*, August 23, https://www.lawfareblog.com/data-brokers-are-advertising-data-us-military-personnel

TAILLAT S., CATTARUZZA A. & DANET D. (2018), *Cyberdéfense. Politique de l'espace numérique*, Paris, Dunod / Armand Colin, Collection U.

TAKAGI K. (2022), *New Tech, New Concepts: China's Plans for AI and Cognitive Warfare*, War on the Rocks, April 13, https://warontherocks.com/2022/04/new-tech-new-concepts-chinas-plans-for-ai-and-cognitive-warfare/