
LES PUBLICATIONS

# Le défi du *big data* et de l'intelligence artificielle pour le monde de la défense – 1e partie

**Par Geoffrey Davril,**

**Avec la participation de Valentyn Baudemont,**

**Membres du comité industrie de défense des Jeunes IHEDN**

LES JEUNES IHEDN

# À PROPOS DE L'ARTICLE

« *Big data* », « intelligence artificielle », « *cloud computing* », « *internet of things* », « transformation digitale » ou encore « *blockchain* » sont les éléments phares du moment dans la sphère numérique. Ces technologies s'implantent et produisent leurs effets dans tous les secteurs d'activité. Elles véhiculent un certain nombre d'avantages et soulèvent plusieurs problématiques : sécurité, gouvernance de la donnée, questions éthiques, transformation des compétences.

Les armées n'échappent évidemment pas à ce « tsunami de données » dont l'expression est empruntée au général Ferlet, directeur du renseignement militaire. Le domaine du numérique se présente comme une nouvelle révolution technologique et industrielle, qui se distingue des précédentes eu égard à sa vitesse de développement. Elle requiert une capacité d'adaptation (dénommée « agilité » dans la sphère numérique) du ministère des armées afin de conserver sa supériorité technique et tactique. La ministre Florence Parly déclarait le 5 avril dernier lors d'un discours portant sur la place accordée à l'intelligence artificielle[14] :

« Alors oui, les armées françaises investissent et investiront dans l'intelligence artificielle, c'est une évidence. Car c'est une technologie stratégique, indispensable pour garantir notre supériorité opérationnelle. »

# À PROPOS DE L'AUTEUR

**Geoffrey Davril** est membre du comité industrie de défense des Jeunes de l'IHEDN.

*Ce texte n'engage que la responsabilité de l'auteur. Les idées ou opinions émises ne peuvent en aucun cas être considérées comme l'expression d'une position officielle de l'association.*

*Crédits de la photo de couverture : état-major des armées.*

# État des lieux

## La donnée, le nouvel or noir

En décembre 2018, le scandale Cambridge Analytica éclate[1]. Les données recueillies par Facebook auraient été utilisées abusivement par des tiers à des fins d'influence électorale aux États-Unis.

Si la nouvelle surprend, ce n'est pourtant plus un secret que nombre de sociétés collectent les données privées émises par leurs utilisateurs et les exploitent à des fins commerciales. Le marché est juteux et fait le succès des start-up de la Silicon Valley. De nombreux capteurs permettent de récupérer de la donnée : lorsque l'on navigue sur un site internet, une application mobile, lorsque l'on se déplace avec son smartphone ou que l'on passe un appel. Ces données, généralement récoltées dans un but économique, sont aussi une aubaine pour les services de renseignement. Produire du renseignement sur une personne ciblée (recueil, analyse et synthèse) peut-être relativement aisé pour un service. En revanche, lorsqu'il devient nécessaire d'analyser une grande masse de données, il est impossible pour une équipe humaine – même de taille importante – de traiter toutes ces informations dans un temps raisonnable ; l'apport des mégadonnées (ou *big data*) devient alors important.

Caractériser cette notion de mégadonnées est essentiel puisqu'elle arrive comme une rupture avec les anciens modèles de collecte et surtout de traitement des données au sein des armées et des forces de l'ordre. La masse des éléments recueillis continuera d'augmenter du fait des innovations techniques et des nouveaux besoins identifiés. Le général Ferlet, directeur du renseignement militaire, en explique les enjeux :

[1] Article *Cambridge Analytica*. Wikipedia

« Aujourd'hui, les capacités toujours plus grandes de ces capteurs, disposant d'un débit toujours plus élevé, nous placent en face d'un « tsunami des données ». Nous sommes submergés par des données dont la masse croît de manière exponentielle. Il ne saurait être question d'y faire face en se contentant seulement de demander des moyens supplémentaires en exploitants ou en analystes. Nous devons au contraire trouver des solutions plus innovantes. »[2]

## Fonctionnement du « traitement massif »

Le *big data* commence par une collecte de données brutes, issues de nombreux capteurs. Ces derniers vont au-delà de la représentation commune que s'en fait le grand public. Un capteur enregistre toute information émise par un humain ou une machine, comme par exemple :

- Un clic sur un lien ;
- Un article de presse publié sur un site internet ;
- Un tweet émis par un utilisateur anonyme ;
- Un dossier santé numérisé ;
- Un mouvement aérien au-dessus d'un territoire ou naval dans un océan ;
- L'ensemble des informations relevées dans l'environnement d'un véhicule, qu'il soit terrien, volant, marin ou sous-marin ;
- Une image satellite.

Consécutivement à la captation de ces données, la série d'analyses qui s'en suit donne une interprétation dont la fiabilité, le délai d'obtention et le degré de précision dépendent de deux facteurs principaux :

- Le ou les algorithmes appliqués aux données ;
- La puissance des machines utilisées pour faire tourner ces algorithmes.

[2] Audition du général Jean-François Ferlet, sur le projet de loi de programmation militaire. Compte-rendu n°52 du 8 mars 2018 de la commission de la défense nationale et des forces armées de l'Assemblée nationale

Les solutions permettant l'analyse des données de masse sont couramment appelées intelligences artificielles (IA). Elles déchaînent les passions, en font rêver certains et en effraient d'autres. Les progrès techniques constamment réalisés les rendent toujours plus performantes. Plusieurs types d'IA sont à distinguer, notamment celui « d'apprentissage par renforcement ». Celui-ci consiste à laisser le système « apprendre » au fil de son utilisation et donc à s'améliorer de façon relativement autonome. Bien que cette méthode ne soit pas adaptée à tous les besoins, elle permet d'obtenir dans certains cas de très bonnes prédictions[3].

Nous sommes encore cependant loin, au moment où ces lignes sont écrites, de voir apparaître la Machine d'Harold Finch[4] ou Terminator[5]. En matière de traitement de données de masse, les IA les plus performantes requièrent de grandes capacités de calcul qui ne sont pas accessibles à toute entité, qu'elle soit individuelle ou collective. Aussi, les questions éthiques associées à ces systèmes se posent dès maintenant, dans le civil comme dans le militaire.

## À l'étranger

En septembre 2019, Edward Snowden a publié ses mémoires. Ce lanceur d'alerte américain a rendu public en 2013 les détails de nombreux programmes de surveillance massive mis en place par la National Security Agency (NSA), dont le programme PRISM[6] chargé de stocker et d'analyser des données numériques extraterritoriales.

Les États-Unis indiquent utiliser l'IA en situation opérationnelle dans le cadre d'actions militaires en Iraq et en Syrie[7]. Depuis les années 2000, l'armée américaine développe et utilise des drones semi-autonomes[8]. Elle collabore en outre avec Israël, un des pionniers de la transformation numérique.[9]

[3] ZAFFAGI Marc, *AlphaZero : l'IA de Google DeepMind devient imbattable aux échecs*. 2017
[4] Article *Person Of Interest*. Wikipedia
[5] Article *Terminator*. Wikipedia
[6] Article *PRISM*. Wikipedia
[7] *Artificial Intelligence and National Security*. Congressional Research Service, 2019
[8] ROTH Marcus, *AI Military Drones and UAVS - Current Applications*. Emerj, 2019
[9] COHEN Sagi, *Startrek, Stargate and the Isreaeli Army's Other AI Projects*. Haaretz, 2019.

En Chine, c'est la reconnaissance faciale de masse qui a fait grand bruit[10]. Des millions de caméras filment en permanence le territoire chinois. Il est ainsi possible d'identifier chaque individu et ses comportements. En corrélant ces données aux informations recueillies sur les réseaux sociaux, la République Populaire a mis en place un système de notation à grande échelle [11]. S'agissant d'armement, les technologies numériques prévalent également. La Chine assume poursuivre ses travaux visant à créer des soldats-robots autonomes[12], même si Terminator n'existe pas encore officiellement.

## En France

Bien qu'elle n'ait pas totalement manqué le virage numérique, la France peut apparaître en retrait par rapport aux États-Unis et à la Chine dans l'évaluation des potentiels de l'IA. Elle dispose des ressources nécessaires s'agissant des compétences humaines (ingénieurs et universitaires) et techniques mais elle a pris du retard. Au sein du ministère des armées, la prise de conscience a eu lieu ces dernières années, notamment en 2016 lorsque la France dût se résoudre à recourir à une solution américaine pour satisfaire les besoins du renseignement français pour gérer les mégadonnées.

En 2017, le Comcyber, commandement de cyberdéfense, est créé[13]. Il rassemble sous une entité interarmées les forces cyber de défense et d'offensive. Puis en 2018, c'est au tour de la direction générale du numérique (couramment abrégée DGNum[14]) de naître. Elle a pour objet d'assurer la transformation numérique de l'ensemble du ministère des armées. Enfin, créée en 2018 également, l'Agence d'innovation défense[15] doit jouer un rôle d'accélérateur de l'innovation pour la défense, numérique ou non. La direction

[10] GUEGJ Léa, *La Chine distribue des bons et des mauvais points à ses citoyens.* France Inter, 2019.
[11] *Chine tout est sous contrôle,* Envoyé Spécial, France 2, 2019.
[12] CHAN Melissa K., *China and the U.S Are fighting a Major Battle Over Killer Robots and the Future of AI*. Time, 2019.
[13] Article *Commandement de cyberdéfense*, Wikipedia.
[14] *Direction générale du numérique, chef d'orchestre de la transformation numérique du ministère*. Ministère des armées, 2018.
[15] *La DGA créé une nouvelle unité spécifiquement dédiée aux systèmes numériques*. Ministère des armées, 2019.

générale de l'armement (DGA) bénéficie avec cette agence d'un organisme plus réactif et plus proche de l'agilité souvent requise pour les projets numériques.

Au-delà de la création de plusieurs entités, l'ambition française se manifeste également sur le plan financier puisque la loi de programmation militaire (LPM) 2019-2025[16] met l'accent sur la transformation numérique des armées :

« Cet effort important, sera consacré à hauteur de 50 % aux domaines du renseignement, de la cyberdéfense et du numérique [...]. De même, la transformation numérique sera un vecteur à privilégier. »

Le ministère a annoncé un budget annuel de cent millions d'euros[17] consacré à l'intelligence artificielle et a recruté de nombreux experts du domaine. Ce montant n'inclut pas les dépenses liées aux programmes d'armement du futur qui intégreront par nécessité une forme plus ou moins avancée de systèmes intelligents. La LPM promet par ailleurs une augmentation du budget R&D lui faisant atteindre le milliard d'euros.

Dans cette première partie, nous avons présenté les technologies que sont le *big data* et l'IA, porteuses d'enjeux clefs du secteur numérique et notamment pour les ministères des armées et de l'intérieur. Nous abordons dans une seconde partie les stratégies adoptées par la base industrielle et technologique de défense afin d'investir ces domaines clefs pour l'avenir des programmes numériques régaliens.





[16] Loi de programmation militaire 2019-2025. Legifrance, 2018.
[17] Discours de Florence Parly, ministre des armées, sur l'intelligence artificielle et la défense. Ministère des armées, 2019.

publication@jeunes-ihedn.org