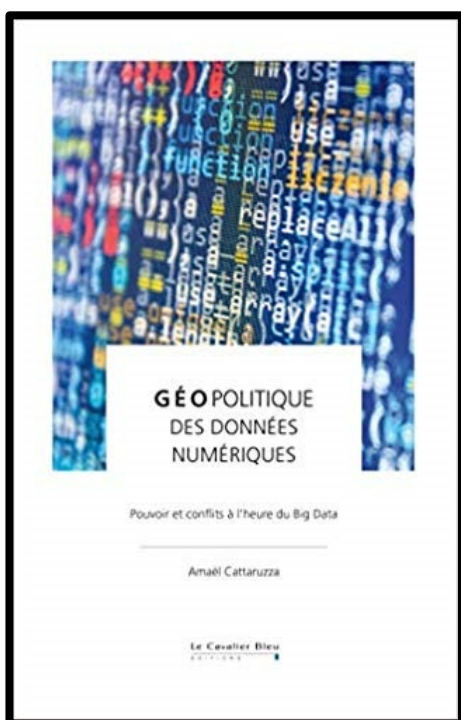


GÉOPOLITIQUE DES DONNÉES. POUVOIR ET CONFLITS A L'HEURE DU BIG DATA

Amaël Cattaruzza, 2019. Paris : Editions Le Cavalier Bleu, 174 p.

Nicolas Bouchaïb, Louis Celhay, Nicolas Bourvic



Amaël Cattaruzza est docteur en géographie et Maître de Conférences à l'École Spéciale Militaire de Saint-Cyr et au Centre de recherche des Écoles de Coëtquidan.

Spécialiste de la géopolitique, il traite plus particulièrement des questions autour des frontières et des Balkans, auxquelles il a consacré plusieurs articles et ouvrages. Après avoir participé en 2013 au colloque organisé à Rennes par l'École des Transmissions de l'Armée de Terre sur « La Cybersécurité : un enjeu mondial, une priorité nationale, des réponses régionales », il décide de se pencher sur le thème du cyberspace. Ses recherches donneront lieu à une publication chez Economica, « La Cyberdéfense. Quel territoire ? Quel droit ? », co-dirigé par Didier Danet, spécialiste de la cyberdéfense.

En 2019, il publie son livre « Géopolitique des données numériques : Pouvoir et conflit à l'heure du Big Data » qui sera, quelques mois plus tard, primé à l'European Cyber Week 2019 (ECW). Dans ce dernier, Amaël Cattaruzza s'emploie à démontrer les modifications géopolitiques entraînées par les données numériques, et ce, en déterminant les spécificités liées aux notions de pouvoirs et de frontières. L'auteur explique également la manière dont la géopolitique, en tant que discipline académique, est bousculée par le cyberspace.

Cet ouvrage est constitué de trois parties : dans un premier temps, l'auteur définit et présente la notion de data, ce qui établit le point de départ de ce travail. Ensuite, il est question de voir comment les données sont « territorialisées » et en quoi cela fait émerger des enjeux à la fois nouveaux mais tout de même transposables à la géopolitique classique. Dans un troisième temps, Amaël Cattaruzza présente une série d'études de cas qui mettent la géopolitique à l'épreuve de la donnée : chaque étude de cas est un exemple d'une situation concrète où la data joue un rôle important dans le dénouement d'un enjeu d'ordre géopolitique.

DEFINIR LA GEOPOLITIQUE DES DONNEES NUMERIQUES

Pour débiter ce travail, Amaël Cattaruzza commence par décrire la notion de donnée, qui constitue le cœur de son ouvrage.

Pour commencer, l'auteur évoque les trois « couches » dont le cyberspace serait constitué et qui permettent de comprendre où résident concrètement les enjeux de la donnée moderne. De ce point de vue, Internet est constitué :

- Tout d'abord de la couche matérielle, qui est l'aspect « physique » d'internet, constitué d'ordinateurs, de *datacenters*, de câbles sous-marins et de serveurs.
- De la couche logicielle, qui est l'aspect « codé » d'internet et qui permet aux ordinateurs de communiquer entre eux.
- Enfin, de la couche sémantique, qui est l'aspect « contenu » d'internet et qui représente l'ensemble des messages et informations contenues sur internet.

À partir de cette structure du web, l'enjeu est alors de savoir ce que l'on appelle « donnée » de nos jours. Ici, la notion de « donnée » fait référence à un processus « d'adaptation et de transcription du réel ». Ces données peuvent être de nature quantitative, c'est-à-dire chiffrée (âge, revenus, nombre d'utilisateurs d'un site...), ou de nature qualitative, c'est à dire modales (nom, adresse, genre...).

Par exemple, sont des datas les informations demandées lors de l'inscription sur un site internet, ou encore les données recueillies de manière non intentionnelle (cookies, fichiers temporaires...), aussi appelées « *exhaust data* ».

C'est donc la donnée, résultante de cette mise en variable, qui est utilisée pour réaliser les calculs relatifs au big data. En reprenant les travaux de Dominique Cardon, l'auteur explique que ces calculs peuvent être de quatre types :

- Au-dessus du web, ces calculs permettent de lier et hiérarchiser les pages internet entre elles. Ces calculs sont à la base du principe de référencement naturel.

- En dessous du web, ce sont des calculs qui transcrivent en données quantitatives les comportements individuels des internautes, afin de prédire leurs futurs agissements.
- Dans le web, ces calculs réalisés autour de l'appréciation de chaque internaute (avis, likes, retweets...), quantifient les réputations et affinités.
- À côté du web, ils permettent, en vérifiant le nombre de clics sur la toile, d'identifier des « publics », c'est à dire des « profils d'utilisateurs ».

Ce qui fait la particularité ainsi que la force du Big Data, n'est pas tant son aspect « Big » dans le sens de l'énorme quantité de données qu'il représente, mais plus spécialement la capacité des machines informatiques à réaliser des calculs précis et rapides avec cet amas de données. Cette puissance de calcul inédite ouvre la porte à de nouveaux travaux de prédictions basés sur les corrélations entre différentes données.

Aujourd'hui, selon certains observateurs cités par Amaël Cattaruzza, avec la puissance de calcul de nos machines et la densité du Big Data, nous devrions bientôt être capables de calculer, corrélérer, prédire puis résoudre des problèmes sociaux ou géopolitiques seulement grâce au traitement automatique des données numériques.

Ce constat amène l'auteur à se poser une question très forte : la perspective d'une telle avancée ne prédit elle pas l'obsolescence programmée des sciences humaines ?

Pour répondre à cette question, l'auteur cite Kitchin : « *Les données n'existent pas indépendamment des idées, des instruments, des pratiques, des contextes et des connaissances utilisés pour les générer, les traiter et les analyser.* » (p.49).

Cette citation est fondamentale pour Amaël Cattaruzza, puisqu'elle montre que la donnée ne peut pas être appréhendée seulement d'un point de vue technique. En effet, à l'instar de Kitchin, l'auteur pense que l'utilisation mathématique du Big Data s'accompagne d'une réflexion humaine qui se doit d'être critique. La donnée, du fait de sa nature, est une perte d'information en soi, car la transposition du réel en données chiffrées apporte une vision imparfaite de l'objet étudié.

Cette notion de réflexion critique vis-à-vis de la donnée est primordiale, puisqu'elle fait naître la problématique de la géopolitique des données, fil rouge du reste de l'ouvrage. En effet, en plaçant la décision humaine au cœur de la dynamique de création des données, la question est ici de comprendre quels sont les jeux de pouvoir à l'œuvre pour contrôler les opportunités qu'apportent ces dernières.

SUPREMATIE DES ÉTATS-UNIS VERSUS LE RESTE DU MONDE

D'après l'auteur, le contexte géopolitique actuel est marqué par la suprématie exercée par les États-Unis.

Elle s'exerce au niveau des trois couches du cyber espace : physique (infrastructures nationales de stockages), logique (industrie numérique nationale) et sémantique (contenus nationaux). Les gouvernements craignent la suprématie des États-Unis, car elle menace leur souveraineté. Pour s'en défendre, plusieurs États commencent à mettre en œuvre des stratégies de « territorialisation » de leurs données. Ces stratégies passent, notamment, par une territorialisation des *datacenters*, faisant ainsi émerger une nouvelle géopolitique des données.

Amaël Cattaruzza donne, par la suite, la définition des *datacenters* : « *Les datacenters sont des centres de stockage et de traitement des données constitués de milliers de serveurs et de commutateurs réseau connectés hiérarchiquement. Ces infrastructures sont à la base des avancées technologiques qui ont permis d'exploiter des données massives, autrefois difficile à traiter.* » (p.71).

C'est cette centralisation des données qui est cœur de l'avancée technologique. Cette dernière a un impact géopolitique d'ampleur, car son stockage doit assurer l'intégrité et la confidentialité des données d'un pays. L'éthique et la confidentialité sont indispensables pour le traitement des données, cependant, peu de pays et d'organisations sont actuellement en mesure de lutter contre la suprématie des États-Unis.

En effet, le PATRIOT ACT, voté en 2001 après les attentats du 11 septembre, permet, par une loi votée aux États-Unis, de contrôler de manière légale les données des individus entrant dans le pays, afin de lutter contre le terrorisme. Mais un abus de cette loi a poussé les américains du nord à utiliser cette dernière à des fins d'espionnage.

En ce qui concerne les infrastructures physiques, la puissance des États-Unis est telle que, sur une carte représentant l'organisation des câbles au niveau mondial, ces derniers apparaissent comme un nœud incontournable pour les flux mondiaux des données : « *97% des échanges de données entre l'Europe et l'Asie transitent par le territoire américain.* » (p. 85).

C'est dans cette optique que les GAFAM (Google, Apple, Facebook, Amazon, Microsoft) ont décidé d'investir en masse dans le marché câblé. Facebook et Microsoft ont payé en 2017, un câble transatlantique allant de la Virginie aux États-Unis à Bilbao en Espagne,

pouvant collecter de nombreuses données entre les deux pays et ainsi asseoir leur suprématie au niveau géopolitique.

Le lanceur d'alerte Edward Snowden, cité par l'auteur, est une des premières personnes à avoir dénoncé la suprématie des États-Unis. En 2013, ce dernier a alerté les autorités de l'espionnage de la population que pouvait effectuer son pays. Cette surveillance de masse a dévoilé, par la suite, les enjeux souterrains de la territorialisation des données.

C'est à partir de l'exemple Snowden, que certains pays ont mis en place des politiques de territorialisation afin de contrôler plus facilement leurs données. L'auteur précise que : « *In fine, cette dynamique de territorialisation à différentes échelles, locale, nationale, régionale, entraîne des rapports de rivalités et de pouvoirs dans le cyberspace, mettant en jeu une large gamme d'acteurs.* » (p.69).

La Russie a décidé, dans cette volonté de régionalisme, de construire des mégadatacenters dans la province de la Sibérie. Cette zone de froid importante du globe permet en effet de faire refroidir les datacenters et, par conséquent, de réduire le coût d'achat de refroidisseurs de machines pour le pays. De plus, le fait de conserver les données personnelles de sa population sur son propre territoire aide à réduire l'influence des États-Unis et la dépendance aux GAFAM.

La localisation de ces infrastructures physiques est un élément majeur dans cette géopolitique des données, car ces infrastructures nécessitent d'être à proximité des points de raccordement, les câbles terrestres ou sous-marins permettant la circulation physique des données. D'après l'auteur : « *Ces enjeux sont d'autant plus importants que 95% du transit des télécommunications et des données Internet mondiales passent à un moment donné ou un autre par des câbles sous-marins.* » (p.85).

LA CYBER SURVEILLANCE DE L'EVERYWARE

Notre rapport à l'espace physique et aux données a été reconfiguré au fil des années par l'extension du numérique dans la totalité de nos activités et occupations.

Cette généralisation nous a amené à ce que Cattaruzza nomme l'époque du « everywhere » ou le « tout logiciel » : « *Que ce soit au travail, en déplacement ou chez soi, l'homme moderne est entouré de code qu'il active* » (p.133).

En effet, les géographes Rob Kitchin et Martin Dodge mettent en avant la notion de « code/espaces » dans leur livre publié en 2011 pour montrer qu'un grand nombre de lieux sont aujourd'hui dépendants du codage issu des données pour fonctionner.

Pour illustrer cette notion, l'auteur emploie l'exemple du voyageur dans l'aéroport. La pertinence de cet exemple est qu'avant de s'installer à la place qui lui est attribuée, le passager devra passer plusieurs zones/espaces qui activeront chacun un ensemble de codes :

1. L'enregistrement : Passenger Name Record = données personnelles, itinéraires, information pour contacter au moins un des participants du voyage, et les informations de la carte de vol. Dans certains cas, d'autres informations comme les réservations d'hôtels, de voitures.
2. Dépose-bagages : type de bagage, poids du bagage, etc.
3. Zones d'embarquement : numéro de la porte d'embarquement, numéro de l'avion, etc.
4. Embarquement : la place attribuée, le type de repas choisi, etc.

Ainsi, à chaque voyageur est donc associé un double, une identité numérique qui lui permettra de passer les différentes séquences de code/espace. Avec l'*everyware*, ce dédoublement se présente sur l'ensemble du monde avec des données associées aux territoires, aux populations ou encore aux marchandises. Le code n'est pas neutre : les informations invisibles à l'œil nu deviennent alors un nouveau type de pouvoir disponible à ceux qui savent y accéder.

Si le contrôle peut être mis à bon escient pour le service d'un passager, à une autre échelle il peut devenir une véritable menace. Le cyberspace s'étant immiscé dans l'ensemble des secteurs, celui de la guerre n'a pas fait exception.

Les instruments de datafication et de transmission des données représentent un atout stratégique mais sont ambigus, car peuvent potentiellement être captés par l'ennemi.

En 2008, le groupe terroriste qui avait perpétré des attaques à Lashkar-e-Taiba du 26 au 29 novembre à Mumbai, avait su profiter d'informations accessibles sur Internet provoquant la mort de 166 personnes et 300 blessés. Grâce aux informations diffusées sur les réseaux sociaux, les terroristes pouvaient s'informer en temps réel sur les mouvements des forces de l'ordre ou sur les décisions opérationnelles des dirigeants, leur permettant de s'adapter quasi-immédiatement. De la même manière, les terroristes pouvaient également dresser le profil des otages qu'ils détenaient.

Cette exploration des données illustre alors le rôle décisif qu'entraîne l'accessibilité aux informations par les deux camps dans ce nouveau type de guerre.

CONCLUSION

L'ouvrage que nous propose Amaël Cattaruzza est riche d'enseignements grâce à ses analyses et ses conclusions qui nous fournissent un large support pour construire nos connaissances sur le sujet. Pour les lecteurs voulant s'initier aux notions de cyberspace, de cybergéographie, « Géopolitique des données numériques : Pouvoir et conflit à l'heure du Big Data » va devenir un livre référence dans son domaine.

L'approche est tout à fait éclairante sur tous les enjeux territoriaux apparus par l'omniprésence de l'Internet et l'importance accrue de la circulation de l'information, ce qui montre le travail d'expertise fourni par l'auteur lors de la rédaction de ce livre.

#CulturesNum est un programme réalisé par les étudiants du **Master Communication des Organisations de l'Université Bordeaux Montaigne** sous la direction d'**Aurélie Laborde**, en collaboration avec **UNITEC**. Depuis 2016, des ouvrages récents sur la société numérique sont synthétisés pour mieux appréhender les questions de fond pour notre société : big data, smart-cities, post humanisme, avenir du travail et de la consommation, etc...