



CLASSIQUES
GARNIER

KHATCHATOUROV (Armen), « Big Data entre l'archive et le diagramme », *Études digitales*, n° 2, 2016 – 2, *Le gouvernement des données*, p. 67-87

DOI : [10.15122/isbn.978-2-406-07064-1.p.0067](https://doi.org/10.15122/isbn.978-2-406-07064-1.p.0067)

La diffusion ou la divulgation de ce document et de son contenu via Internet ou tout autre moyen de communication ne sont pas autorisées hormis dans un cadre privé.

© 2017. Classiques Garnier, Paris.
Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous les pays.

KHATCHATOUROV (Armen), « Big Data entre l'archive et le diagramme »

RÉSUMÉ – Au sein des Big Data, nous distinguons deux tendances qui permettent de décrire les nouveaux “diagrammes” du pouvoir, les modes de gouvernementalité qui leur correspondent, ainsi que leurs relations complexes avec les savoirs. Cette analyse, convoquant le paradigme cybernétique, nous amène à formuler cette hypothèse : il en va désormais non pas simplement d'une nouvelle forme historique de la relation entre les savoirs et le pouvoir mais de l'auto-actualisation indéfinie de ce dernier.

ABSTRACT – Within Big Data, we distinguish two trends that allow us to describe new “diagrams” of power, the modes of governmentality that correspond to them, and their complex relations with knowledge. This analysis, drawing upon the cybernetic paradigm, leads us to formulate this hypothesis: from now on, what is in question is not simply a new historical form of the relationship between knowledges and power, but the indefinite self-actualization of the latter.

BIG DATA ENTRE L'ARCHIVE ET LE DIAGRAMME

DE LA PRÉSUPPOSITION RÉCIPROQUE

On postule aujourd'hui que nos savoirs dépendent de plus en plus des « données » obtenus à l'aide de Big Data. L'approche critique de ce champ, d'abord structurée par des enjeux épistémologiques, thématise la manière dont ce nouveau paradigme redéfinit la constitution des savoirs. Après une brève restitution de ce débat dans le but d'en dégager quelques éléments distinctifs et problématiques, nous allons, dans ce texte, tenter de le compléter par une approche dont l'aspiration est de dépasser les problèmes strictement épistémologiques pour s'acheminer vers ceux de l'exercice du pouvoir, en amont de la constitution des savoirs particuliers. Pour ce faire, nous tenterons de dégager le *diagramme* d'aujourd'hui, dans le sens où ce terme est employé par Gilles Deleuze, à savoir une représentation, une exposition « des rapports de force qui constituent le pouvoir ».

Dans son livre consacré à la pensée de Michel Foucault, Deleuze fait la distinction entre l'archive et le diagramme. L'archive est comprise comme le savoir constitué et le diagramme comme les rapports de forces qui sous-tendent la constitution de tel ou tel savoir. Inversement, tout pouvoir nécessite des savoirs constitués, s'appuie sur eux : sans eux, il « resterait ponctuel, fluent, évanouissant, instable et ne pourrait ni se conserver ni se reproduire¹ ». Mais le rapport entre les deux, s'il est de *présupposition réciproque*, n'est pas symétrique : il y a, selon Deleuze, un primat du pouvoir sur le savoir dans le sens où celui-ci rend tout savoir à la fois possible, à travers sa constitution, *et* problématique, en

1 Gilles Deleuze, 'Foucault – Le Pouvoir, Cours 13', 1986 <http://www2.univ-paris8.fr/deleuze/article.php3?id_article=453>.

lui conférant un caractère provisoire. Soit une opération de recherche lexicologique sur un corpus par exemple : « si l'on cherche à déterminer un corpus de phrases et de textes pour en extraire des énoncés, on ne peut le faire qu'en assignant les foyers de pouvoir (et de résistance) dont dépend ce corpus² ».

La relation archive – diagramme semble à première vue analogue au schématisme kantien :

Le diagrammatisme de Foucault, c'est-à-dire la présentation des purs rapports de forces ou l'émission des pures singularités, est donc l'analogue du schématisme kantien : c'est lui qui assure la relation d'où le savoir découle, entre les deux formes irréductibles de spontanéité et de réceptivité³.

Mais il n'y a ici point de téléologie : là où, chez Kant, le scientifique devait invariablement faire comme si la constitution du savoir absolu était possible, Foucault prend non seulement la mesure du caractère fini de tout savoir constitué mais dépeint la situation où le processus de sa constitution est lui-même un produit provisoire de rapport de forces, comme l'illustre par exemple la dépendance des savoirs médicaux ou pénitentiaires de ce diagramme particulier qu'est le panoptique. Ainsi, le pouvoir « produit du vrai comme problème⁴ », et sa relation à l'archive se joue précisément dans l'historicité des « régimes de vérité ».

Chaque époque, chaque formation historique, donne à voir ses propres diagrammes. On le sait, Foucault a consacré l'essentiel de son travail à l'analyse de cette époque particulière où se jouait la conjonction, la confrontation, entre deux diagrammes : celle des disciplines et de l'anatomo-politique d'un côté, et celle de sécurité et de la biopolitique de l'autre. Plus près de nous, le néolibéralisme⁵ est une formation nouvelle qui rejoue cette confrontation sous une forme de « société de contrôle » dont on trouve les premiers éléments analytiques chez Foucault lui-même, et dont les modalités concrètes sont encore à déterminer.

2 Gilles Deleuze, *Foucault* (Les Éditions de Minuit, 1986).

3 Deleuze, *Foucault*.

4 Deleuze, *Foucault*.

5 Il faut préciser dès maintenant que nous donnons à ce terme le sens précis qu'il a chez Foucault et non ses multiples sens courants, que le « sens courant » soit entendu dans son sens initial de la deuxième moitié du XIX^e, ou dans le sens de la réaction anti-keynésienne des années 1980.

Si notre intuition est que le régime néolibéral lui-même est en train de subir des transformations profondes, nous n'abordons pas ici cette question frontalement mais à travers une tentative d'en dégager les diagrammes nouveaux, au moins ceux qui nous semblent à l'œuvre dans le domaine de Big Data. Et, comme nous essayerons de le montrer dans la conclusion, la particularité de ce domaine ne réside peut-être pas tant dans le passage vers un diagramme particulier et vers les savoirs particuliers qui l'actualisent que dans le changement de type de rapport, de la nature même de la présupposition réciproque, qui unit l'archive et le diagramme.

BIG DATA ENTRE ENJEUX ÉPISTÉMOLOGIQUES ET SOCIÉTAUX

Quelle est la nouveauté, à l'ère de Big Data, dans les rapports entre le pouvoir et les savoirs ? Tentons d'en donner une première approche à travers une analyse d'ordre épistémologique.

L'émergence de Big Data correspond d'abord à de nouveaux modes de récolte et de traitement des données, voire à un nouveau mode de rapport aux phénomènes décrits, que ces phénomènes relèvent des sciences de la nature ou des sciences humaines et sociales. Le domaine du Big Data bénéficie aujourd'hui d'une attention croissante et fait objet d'investissements massifs tant scientifiques qu'industriels. Cependant, après une première phase d'euphorie, illustrée par l'article fondateur et sans doute volontairement polémique de Chris Anderson⁶, l'heure est venue d'aborder le Big Data avec plus de recul, en examinant en détail aussi bien les prémises que les conséquences de sa généralisation.

Le Big Data recouvre des réalités très différentes. Il est porteur de confusions terminologiques qui non seulement n'aident pas à la compréhension des enjeux adjacents mais, nous semble-t-il, alimentent des discours qui profitent de cette confusion. Ainsi, dans un souci de

6 Chris Anderson, 'The End of Theory : The Data Deluge Makes the Scientific Method Obsolete', *Wired*, 2008, p. 1-2 <<http://dx.doi.org/10.1016/j.ecolmodel.2009.09.008>>. Paru également dans *Edge – The Third Culture*.

clarification conceptuelle, Francis Diebold⁷ distingue entre Big Data le *terme*, Big Data la *discipline* et Big Data le *phénomène*, acceptions auxquelles nous ajoutons Big Data le *discours*. Comme tout phénomène technique, il n'est compréhensible qu'en conjonction avec le discours qui le porte ; comme nous le verrons tout à l'heure, une analyse de ces discours semble nous rapprocher déjà des enjeux de pouvoir.

Pour caractériser le Big Data, on emploie souvent la métaphore des 4 V :

- Le Volume (la grande quantité des données qui nécessite souvent des capacités de stockage importantes, voire le stockage repartit) ;
- La Variété (la variété des formats et des sources, par exemple lorsqu'un traitement porte à la fois sur les fichiers audio et texte) ;
- La Vitesse (le « temps réel » du traitement) ;
- La Vérité (la mise en place des méthodes spécifiques pour gérer l'imprécision des données, leur intégrité, leur adéquation au phénomène étudié).

S'appuyant sur des techniques spécifiques, l'analyse de ces données donnerait lieu à la découverte des patterns et des corrélations qui ne sont pas présumés ou même soupçonnés à l'avance. Selon le « discours Big Data », nous passerions ainsi du mode de raisonnement déductif au mode inductif, mode qui ne s'appuie pas sur des théories préalables et qui est alors présenté comme le plus neutre et objectif possible en suivant un argument de type « puisque les données nous le disent... ». On serait alors en présence d'un nouveau paradigme s'appuyant sur la capacité scientifique et technique de croisement des sources et de leur traitement. Dans cette nouvelle manière de voir le réel, on pourrait se passer des hypothèses à tester, on serait en mesure de procéder à la découverte du réel « tel qu'il est ». Somme toute, on serait en présence d'un nouvel empirisme débarrassé de ses inconvénients puisque s'appuyant sur des procédures techniques bien précises. Le Tableau 1 illustre ce changement supposé de paradigme.

⁷ Francis Diebold, *A Personal Perspective on the Origin(s) and Development of 'Big Data' : The Phenomenon, the Term, and the Discipline*, 2012.

Paradigm	Nature	Form	When
First	Experimental science	Empiricism ; describing natural phenomena	pre-Renaissance
Second	Theoretical science	Modeling and generalization	pre-computers
Third	Computational science	Simulation of complex phenomena	pre-Big Data
Fourth	Exploratory science	Data-intensive ; statistical exploration and data mining	Now

TAB. 1 – L'hypothèse du Big Data comme le 4^e paradigme⁸.

Cependant, on est en droit de se demander si un tel discours ne pêche pas par un enthousiasme aveugle, voire ne cache pas d'autres enjeux.

La critique épistémologique de Big Data se situe alors sur le terrain de ce que plusieurs travaux identifient comme le nouveau mode de génération des hypothèses. Là où les sciences d'avant le Big Data auraient procédé sur un mode déductif, elles seraient, à l'heure actuelle, en train d'explorer le mode inductif. Ce postulat semble en lui-même hautement problématique ainsi que nous allons maintenant expliciter⁹.

Tout d'abord, en quoi consiste la nouveauté de Big Data ? Lorsqu'on regarde de plus près, on se rend compte que la plupart des approches scientifiques sont en réalité anciennes. Presque tous les traits distinctifs de Big Data, tant au niveau des données que des méthodes, puisent leurs racines dans des approches qui datent au moins d'une vingtaine d'années. Pour ne prendre que quelques exemples : la variété des formats est traitée depuis longtemps par les méthodes dites « sensor fusion » ; la présence de l'approche inductive a sa place dans la science depuis toujours comme

8 R. Kitchin, 'Big Data, New Epistemologies and Paradigm Shifts', *Big Data & Society*, 1.1 (2014) <<http://dx.doi.org/10.1177/2053951714528481>>.

9 Des travaux critiques allant dans ce sens sont de plus en plus nombreux. On consultera notamment à cet effet : *Raw Data Is an Oxymoron*, ed. by Lisa Gitelman (MIT Press, 2013), le numéro spécial de *First Monday Making Data — Big Data and beyond* (First Monday, vol. 18, number 10, 2013), XIX., le numéro spécial de *International Journal of Communication Critiquing Big Data : Politics, Ethics, Epistemology | Special Section* (International Journal of Communication, Vol. 8, 2014), ainsi que Antony Bryant and Uzma Raja, 'In the Realm of Big Data ...', *First Monday*, 19.2 (2014) <<http://dx.doi.org/10.5210/fm.v19i2.4991>> et Kitchin, *op. cit.*

le soulignait Charles Sanders Peirce il y a plus d'un siècle ; la recherche « tirée par les données » n'est pas nouvelle en soi puisque les Big Data d'aujourd'hui se nourrissent des approches probabilistes existantes. Les seuls éléments qui échappent clairement à l'enracinement ancien, nous y reviendrons, sont le volume en tant que tel et la question du passage de l'échantillon fini à la population complète ($n=all$).

On le sait bien, comme l'illustre par exemple l'histoire des sciences cognitives, que la mise en commun des approches différentes peut avoir des vertus propres, leur hybridation mutuelle conduisant à de nouvelles manières fructueuses d'aborder les questions anciennes. Néanmoins, il faut, à notre sens, nuancer ces formules chocs autour du changement de paradigme.

Ensuite, les données parlent-elles d'elles-mêmes ? Il y a bien un nouveau mode de générer les hypothèses, mais ces hypothèses restent guidées par les connaissances existantes et les présupposés des chercheurs. Les données et les résultats sont plus que jamais construits, que cela soit par le choix des données à récolter, le choix des modèles ou même le choix de tel ou tel logiciel de traitement.

On a alors l'impression que la nouveauté de Big Data réside plutôt dans une combinaison spécifique des méthodes scientifiques et des enjeux épistémologiques : comment construit-on les modèles ? Quels agendas guident ces choix ? Mais plutôt que de s'intéresser à la construction sociale de la science, telle que Bruno Latour par exemple l'a thématisée en proposant en particulier de remplacer le terme de « données » par celui d'« obtenues¹⁰ », les discours Big Data s'emploient à passer sous silence cette construction, là où, paradoxalement, elle devrait être la plus apparente puisque toute action de récolte ou de traitement des données est potentiellement traçable et pourrait faire l'objet d'une analyse d'ordre épistémologique.

On peut ici entrevoir l'articulation des enjeux épistémologiques et des enjeux de pouvoir en prenant au sérieux Big Data le *discours* que nous évoquions tout à l'heure. Par exemple, certains auteurs distinguent dans ce discours deux métaphores principales et solidaires qui décrivent toutes deux la supposée neutralité des données en les assimilant soit aux forces de la nature soit aux ressources à consommer : « Big Data

10 Bruno Latour, *L'espoir De Pandore : Pour Une Version Réaliste De L'activité Scientifique* (Paris : La Découverte, 2001).

is a force of nature to be controlled, an uniform mass », et « Big Data is nourishment / fuel to be consumed, new oil¹¹ ». C'est en prônant la nouveauté des méthodes sans en faire une généalogie, c'est en organisant la promotion de la maîtrise et de la richesse que le pouvoir établit l'impératif pour toute action, y compris politique ou sociale, de passer par les données – et s'établit ainsi comme l'instance qui en détermine la circulation. Or, si l'on tient à poursuivre la métaphore des matières premières ou des forces de la nature pour décrire Big Data, c'est moins du côté du pétrole que du côté de l'atome et de ses conséquences ambiguës qu'il faudrait sans doute chercher. Le risque encouru est que, derrière la fascination pour les données et les discours annonçant des changements épistémologiques, on fasse croire que les outils et les méthodologies sont neutres et que les données « parlent d'elles-mêmes ». Or, il n'y a pas de données « brutes¹² ». Les données sont plus que jamais construites selon les choix industriels et politiques qui orientent non seulement la recherche mais la société dans son ensemble, choix qui risquent alors d'être passés sous silence.

Notre hypothèse est cependant que les enjeux de Big Data ne se résument ni à des problèmes strictement épistémologiques, si l'on entend par là le mode de génération des hypothèses, ni à la présence évidente des discours « marketing », y compris politiques et scientifiques, qui chercheraient à mettre en avant la nouveauté supposée des méthodes en question. Il faut également, nous semble-t-il, aller au-delà de la critique même des métaphores des « forces de la nature » ou des « ressources naturelles » et, partant, au-delà d'une certaine compréhension heideggérienne de la technique comme Arraisonement (*Gestell*) – ou du moins tenter d'en décrire des formes nouvelles¹³. Pour ce faire, nous proposons de déceler dans le phénomène actuel de Big Data deux diagrammes, deux tendances en partie concurrentes, que nous appellerons *l'exhaustivité mémorielle* et *l'immédiateté*.

11 Cornelius Puschmann and Jean Burgess, 'Metaphors of Big Data', *International Journal of Communication*, 8 (2014), 1690–1709.

12 *Raw Data Is an Oxymoron*, ed. by Lisa Gitelman (MIT Press, 2013)

13 Il est en effet frappant de voir que l'analyse critique de Big Data *le discours* que nous avons évoqué reproduit le versant le plus connu de la critique heideggérienne de la technique moderne : le Rhin est à la fois maîtrisé, emprisonné et fournit de l'énergie disponible. Or, avec Big Data il ne s'agit plus simplement de cela. Cf. Martin Heidegger, 'La Question de La Technique', in *Essais et Conférences* (Paris : Gallimard, 1993, 1954).

DEUX DIAGRAMMES DE BIG DATA

LE DIAGRAMME DE L'EXHAUSTIVITÉ MÉMORIELLE :
BIG DATA WAREHOUSE

Dans sa variante la plus connue et la plus ancienne, pour laquelle nous employons ici le terme de *Big Data Warehouse* (entrepôt de données), le Big Data consiste en une accumulation la plus exhaustive possible des données. L'exhaustivité concerne ici aussi bien la couverture d'une population donnée ($n=all$) que la description, réelle ou fantasmée, de chaque individu¹⁴. À notre sens, ce mouvement est solidaire de l'impératif de visibilité absolue, les discours sur la science et sur la société étant ici analogues. En effet, du côté de Big Data « scientifique », on prône l'accès direct au réel à partir d'un relevé exhaustif par les données pouvant être réalisé à la demande ; du côté de Big Data « financier », « aide à la décision » ou « surveillance », on aspire à décrire les individus de manière exhaustive en réduisant chacun d'eux à un jeu complet d'attributs de plus en plus fins et, dans le cas particulier des individus humains, on procède à la détermination de l'existence individuelle à partir de ses traces. Dans ce mouvement conjoint, ces deux approches contribuent à un nouveau régime de visibilité qui consiste à transformer le réel en données et à le réduire par là même à ce qui peut être saisi « en données », rendant êtres et objets « à disposition ». Certes, la critique latourienne (qui portent sur la confusion entre les *données* et les *obtenues*) et heideggérienne (qui porte sur la réduction du réel au stock disponible et maîtrisé) sont ici plus que justifiées. De même, le phantasme que nous avons déjà évoqué, celui de disposer d'un accès direct, sans intermédiaire d'hypothèses ou de catégories préalables, est ici plus que critiquable. Mais il nous semble que l'essentiel est ailleurs.

La question, beaucoup plus simple mais qui nous semble plus cruciale et qui nous rapproche du diagramme d'aujourd'hui, est celle du régime des finalités selon lequel fonctionne le *Big Data Warehouse*. On peut en

14 Les termes de « population » et d'« individu » sont ici employés dans le sens que la statistique leur donne.

effet légitimement se demander quelles finalités guident cette récolte acharnée des données. Qu'est-ce que l'on souhaite en faire ? Quelles méthodes sont à employer ?

L'exemple qui est ici sans doute le plus parlant est celui des données à caractère personnel, ce nerf de la guerre économique que les acteurs majeurs du domaine se livrent. Le concept de finalités de traitement apparaît dans les années 1970, lorsque le croisement des données des différentes bases de données gouvernementales risque de conduire à un déséquilibre des pouvoirs et de menacer les libertés individuelles. Ce concept correspond alors à l'injonction de savoir à l'avance la destination de la collecte et du traitement des données. Les données personnelles « sont collectées pour des finalités déterminées, explicites et légitimes et ne sont pas traitées ultérieurement de manière incompatible avec ces finalités¹⁵ ». Parallèlement et de manière liée, apparaissent deux autres concepts : celui des données à caractère personnel (données potentiellement identifiantes) et celui du consentement (comme l'injonction de demander à l'utilisateur son accord pour la collecte).

Ces trois pierres angulaires sont remises en question par le Big Data. En effet :

- Pas de finalité préalable à la collecte et au traitement : on collecte tout ce que l'on peut, tout d'abord parce que les coûts de cette collecte sont de moins au moins élevés, et plus fondamentalement parce que les données peuvent servir à quelque chose plus tard, conformément au postulat épistémologique de recherche de corrélations « insoupçonnées » ;
- Plus de distinction entre les données identifiantes et non-identifiantes : toutes les données sont potentiellement identifiantes, soit par la ré-identification dans les bases existantes, soit par le croisement avec de nouvelles données collectées sans finalité préalable (*cf.* ci-dessus) ;
- Pas de consentement, tout d'abord parce que la collecte se fait, de fait, à l'insu de l'utilisateur, et plus fondamentalement parce qu'il est impossible de demander le consentement pour des finalités non-définies (*cf.* ci-dessus).

15 *Loi Informatique et Libertés* (Loi n° 78-17 du 6 janv. 1978 relative à l'informatique, aux fichiers et aux libertés : JO, 7 janv. 1978).

On assiste alors à une récolte de données guidée par un fantasme d'exhaustivité qui se joue sur deux plans. Le premier est celui de la profondeur temporelle : le rapport au temps s'assimile à une mémoire infinie qui peut en droit (si ce n'est en fait, limité provisoirement par les contraintes techniques de stockage) aller indéfiniment dans le passé. Faut-il insister ici sur la durée de plus en plus grande de conservation des données à caractère personnel permise par les législations successives, sous couvert de lutte antiterroriste ? Le deuxième plan est celui de l'extension : le rapport à la population change radicalement en passant de l'échantillon représentatif à la couverture totale, illustrée pas la formule $n = all^{16}$.

De cette absence de finalités, il convient de prendre la juste mesure. On pourrait d'abord évoquer le fait que, dans la pratique qu'ont aujourd'hui les ingénieurs, il ne s'agit pas tant d'accomplir une finalité que l'on se serait fixée à l'avance que d'être surpris par des retombées plus ou moins insoupçonnées, dans un jeu perpétuel de confrontation avec les algorithmes informatiques¹⁷. Si l'on peut risquer cette formule, il ne s'agit plus d'aspiration à la maîtrise mais à la surprise ; et s'il s'agit encore d'une « mise en réserve », celle-ci réserve l'insoupçonnable.

Plus fondamentalement, n'en va-t-il pas d'une conséquence prévisible de l'époque moderne, voire de son aboutissement ? C'est en effet le propre de la technique en général que d'opérer le déplacement perpétuel des fins, déplacement qui ne peut qu'aboutir à la remise en question de l'idée même de finalité préalable à une action quelle qu'elle soit. Comme le souligne Jean-Luc Nancy, « la technique et toute notre époque [...] nous étalent une sorte d'absence de fin, d'accomplissement, de but, de téléologie¹⁸ ». Le propre de *Big Data Warehouse* serait alors de porter cette

16 C'est peut-être là la radicalisation *et* la fin de la biopolitique à laquelle Foucault assigne encore la fonction de gouverner, non pas les individus ou les corps, mais la vie, la population prise dans sa durée et les événements aléatoires qui s'y produisent. Or, la vie elle-même est de plus en plus comprise comme ce qui se révèle dans les données, et la population comme population exhaustive. Dès lors, s'agit-il encore de la vie-bios, ou de la vie-zôê, « déjà devenue tekhnè » ? (Jean-Luc Nancy, 'Note Sur Le Terme de « Biopolitique »', in *La Création Du Monde Ou La Mondialisation* (Galilée, 2002) p. 140). C'est peut-être là une des différences qui séparent la « biopolitique » libérale de Foucault de notre néolibéralisme tardif, du moins dans son aspect d'aspiration à l'exhaustivité.

17 Jean-Pierre Dupuy, 'Le Problème Théologico-Scientifique et La Responsabilité de La Science', *Le Débat*, 2004, 182.

18 Jean-Luc Nancy, '« Techniques Du Présent », Entretien Avec Benoît Goetz', *Le Portique*, 3 (1999).

absence à son paroxysme. Plus que de l'*Arraisonnement* heideggérien, il s'agit ici d'une remise en cause du régime même de finalités, régime selon lequel l'action humaine est guidée.

Tel est, nous semble-t-il, le premier diagramme de Big Data : une exhaustivité mémorielle dans les deux sens d'extension temporelle et de couverture de la population, sans finalité préétablie.

LE DIAGRAMME DE L'IMMÉDIATÉTÉ : *BIG DATA STREAM*

Le second diagramme semble être guidé par les contraintes de réactivité par rapport aux phénomènes traités et se voit imposer les délais de réponse dits « en temps réel ». Il puise ses racines dans *data stream mining* relativement ancien, puisque datant du début des années 2000. L'approche change ici radicalement : on passe du paradigme « capturer – analyser » à l'analyse à même le flux de données. On parle dès lors de *Big Data Stream* (flux de données).

En langage informatique, le *Big Data Stream* s'oppose radicalement au *Big Data Warehouse* de la même manière que le *stream processing* s'oppose au *batch processing* : le traitement ne repose pas ici sur une certaine reproductibilité, sur une série d'instructions stable appliquée à des corpus de données finis et sur le travail par lots, mais sur l'écoulement indéfini de données, où seuls les résultats de l'analyse sont éventuellement stockés, quand ils ne servent pas à leur tour d'input à d'autres flux et d'autres analyses¹⁹. Il va alors de soi que le phénomène analysé dépend de la grille d'analyse de manière beaucoup plus intime que dans *Big Data Warehouse* où l'on peut revenir à tout moment à l'archive (que ce retour soit réel ou fantasmé)²⁰.

La métaphore utilisée ici est celle de la recherche de l'or, où l'épistémologie et l'économie ne font plus qu'un : si la grille qui sert à « analyser » le flux est trop fine, trop grosse ou ne correspond pas à

19 « Stream Processing Engines (SPEs) are specifically designed to deal with streaming data [...] SPEs perform SQL-style processing on the incoming messages as they fly by, without necessarily storing them. Clearly, to store state when necessary, one can use a conventional SQL database embedded in the system for efficiency. SPEs use specialized primitives and constructs (e.g., time-windows) to express stream-oriented processing logic. » (Michael Stonebraker, Uğur Çetintemel and Stan Zdonik, 'The 8 Requirements of Real-Time Stream Processing', *ACM SIGMOD Record*, 34.4 (2005), 42–47 <<http://dx.doi.org/10.1145/1107499.1107504>>.)

20 « [...] as Heraclitus 2.0 might say, you cannot step twice into the same data stream » P. Dourish, 'Algorithms and Their Others : Algorithmic Culture in Context', *Big Data & Society*, 3.2 (2016) <<http://dx.doi.org/10.1177/2053951716665128>>.

la forme de l'or, on ne « trouve » pas ce que l'on cherche, et on alloue alors les ressources humaines ou de calculs à d'autres flux, à d'autres *data pipelines* plus rentables.

*The analogy here is that working in data streams is much like panning for gold. As the data stream passes by, analysis occurs in parallel that seeks to capture the relevant nuggets of information to best address specific questions or areas of concern*²¹.

En d'autres termes encore, dans *Big Data Stream*, la capacité d'analyse est déportée auprès des capteurs, dont le rôle, désormais, n'est plus de fournir un relevé étalonné des phénomènes, ni de « fabriquer » les phénomènes de la science mais, – interconnectés entre eux et composant ainsi l'Internet des objets (ou plutôt, selon une expression désormais plus appropriée, *Internet of Everything*) –, de constituer le monde lui-même.

*We expect to see everything of material significance on the planet get “sensor-tagged” and report its state or location in real time. This sensorization of the real world will lead to a “green field”²² of novel monitoring and control applications with high-volume and low-latency processing requirements*²³.

Il faut ici en prendre la mesure à l'aide des cas concrets. Prenons l'exemple de *capteur virtuel*, ce paradigme opérationnel issu du domaine de « sensor fusion ». Le capteur virtuel est une définition algorithmique de la mise en correspondance d'éléments hétérogènes :

*Virtual sensors abstract from implementation details of access to sensor data and correspond either to a data stream received directly from sensors or to a data stream derived from other virtual sensors. A virtual sensor can be any kind of data producer, for example, a real sensor, a wireless camera, a desktop computer, a cell phone, or any combination of virtual sensors. A virtual sensor may have any number of input data streams and produces exactly one output data stream based on the input data streams and arbitrary local processing*²⁴.

21 Rob Chimsky, 'Diving into Big Data : Data Lakes vs. Data Streams', *Siliconagle*, 2014 <<http://siliconagle.com/blog/2014/04/17/diving-into-big-data-data-lakes-vs-data-streams/>>.

22 L'expression « green field » est ici manifestement à comprendre dans son acception de la langue des affaires : un nouveau projet, une nouvelle construction sur un champ non-défriché, une création *ex nihilo*.

23 Stonebraker, Çetintemel and Zdonik, *op. cit.*

24 Karl Aberer, Manfred Hauswirth and Ali Salehi, 'Infrastructure for Data Processing in Large-Scale Interconnected Sensor Networks', in *2007 International Conference on Mobile Data Management* (Conference and Custom Publishing, 2007), p. 198–205 <<http://dx.doi.org/10.1109/MDM.2007.36>>.

Ainsi, l'output de *Big Data Stream* est basé à la fois sur les événements « réels », sur les combinaisons de ces capteurs virtuels dont certaines entrées sont elles-mêmes des événements purement logiciels et sur les traitements algorithmiques pratiqués à même ces flux de données. Cet output produit ainsi des événements dits « composés » ou « dérivés ». Si l'on se réfère par exemple à une des variantes de *stream processing* dite CEP, un événement « observé » correspond fondamentalement, non à la « cause » ou à l'input, mais à la conséquence d'un traitement algorithmique.

A special case of stream processing is complex event processing (CEP). CEP terminology refers to data items in input streams as raw events and to data items in output streams as composite (or derived) events. A CEP system uses patterns to inspect sequences of raw events and then generates a composite event for each match, for example, when a stock price first peaks and then dips below a threshold²⁵.

[The] derived streams often do some kind of enrichment, say adding on new attributes not present in the original event. Derived streams require no particular handling. They can be computed using simple programs that directly consume from Kafka and write back derived results or they can be computed using a stream processing system. Regardless which route is taken the output stream is just another Kafka topic so the consumer of the data need not be concerned with the mechanism used to produce it²⁶.

Bien entendu, il ne s'agit pas ici de dénoncer un voile de données posé sur une réalité préalable mais de comprendre un mode de « dévoilement » du monde qui n'est plus simplement celui de la maîtrise des forces de mise à disposition ou de commande. Ce traitement de données, qui a pour caractéristique d'injecter des procédures qui ne peuvent pas être qualifiées de simple analyse, fait advenir en temps réel des patterns d'événements. Le monde, sans profondeur bien qu'allant très loin dans la fouille des données, est ici celui de décisions immédiates et automatisées, basées sur la circulation rhizomatique de données ; celui de décisions pouvant servir à leur tour d'inputs consécutifs dans d'autres traitements, selon un jeu perpétuel d'ajustements des variables entre elles, dans le seul but, lui-même malléable, de prédire et de réguler de comportements.

25 M. Hirzel and others, 'IBM Streams Processing Language : Analyzing Big Data in Motion', *IBM Journal of Research and Development*, 57.3 (2013), 7:1-7 :11 <<http://dx.doi.org/10.1147/JRD.2013.2243535>>.

26 Jey Kreps, 'Putting Apache Kafka To Use : A Practical Guide to Building a Stream Data Platform (Part 2)', *Confluent*, 2015 <<http://www.confluent.io/blog/stream-data-platform-2/>>. *Kafka* est le nom du logiciel open source pour le *stream processing*, utilisé notamment par LinkedIn, Netflix, Spotify et Uber. Nous soulignons.

Ainsi, le phénomène « obtenu » est un résultat de processus de contrôle : contrôle non dans le sens d'un message dirigé de commande, mais dans le sens dont on retrouve l'ébauche, en 1942, dans l'*Anti-Aircraft Predictor* de Norbert Wiener. Rappelons ici que le *AA Predictor* a été conçu pour prédire le comportement du pilote ennemi compris comme un servomécanisme, dont la trajectoire future pourrait être anticipée en fonction de ses automatismes propres, des contraintes mécaniques du vol, de la trajectoire passée dans une fenêtre de temps donnée, etc. Il nous semble que la vraie rupture dont le *AA Predictor* a été porteur n'est pas tant dans la réduction conceptuelle du pilote humain à une machine à feedback, ni dans la considération du système pilote-avion comme un tout fonctionnel. La vraie rupture consiste dans le fait que le canonier lui-même est alors conçu de la même manière que le pilote et devient l'« environnement » pour ce dernier (ce qui, dans peu de temps, se confirmera factuellement par l'avènement de l'équipement embarqué de l'avion et du drone de combat comportant la même capacité d'analyse des patterns des comportements ennemis).

In the airground battle, it was a short step for Wiener and Bigelow to take the pilot-as-servomechanism directly over into the AA gunner-as-servomechanism and thence to the operation of the heart and proprioceptive senses²⁷.

You start with an input-output approach which is suggested by the fact that you have no access to the pilot, you make this enormously successful prediction of where the plane would be, and then you apply that back to your own gunner. By now, the train of generalizations is rolling and you can see then how it gets generalized to other systems. [...] The cybernetic model in some way grew up around an enemy that was symmetrical, which is why you could move easily from the enemy taking evasive action to you taking evasive action or to the gunner taking anti-evasive action, but it was all in some sense a symmetric piece²⁸.

Réversibilité de l'ennemi et de l'allié, réversibilité du phénomène analysé et de l'instance d'analyse, réversibilité de l'environnement et de l'individu ; action sur les règles qui « décrivent » l'environnement²⁹ : tels

27 Peter Galison, 'The Ontology of the Enemy : Norbert Wiener and the Cybernetic Vision', *Critical Inquiry*, 21.1 (1994), 228 <<http://dx.doi.org/10.1086/448747>> p. 264.

28 Peter Galison and Sina Najafi, 'The Ontology of the Enemy : An Interview with Peter Galison', *Cabinet Magazine*, 2003.

29 Nous ne développons pas ici ce point, mais il semble évident que les règles qui décrivent et constituent l'environnement pour le canonier de *AA Predictor* ne relèvent pas simplement de la « physique » des processus mais sont réifiées par la machine ou, pour employer des termes plus contemporains, par le « machine learning ». Comme le note Galison :

sont les traits distinctifs de la conversion cybernétique. Ce paradigme correspond alors à un ajustement mutuel dont personne ne détient l'initiative, paradigme dont la portée dépasse le problème de détection de patterns dans l'exemple de l'avion de combat.

Car il est à ce titre frappant de constater comment Foucault, sans doute très au fait des recherches de son époque dans le domaine de la cybernétique sans qu'il les mentionne explicitement, décrit le mode alors naissant de la gouvernementalité néolibérale et de l'exercice du pouvoir qui lui correspond.

Ce qui apparaît, ce n'est pas du tout l'idéal ou le projet d'une société exhaustivement disciplinaire dans laquelle le réseau légal, enserrant les individus, serait relayé et prolongé de l'intérieur par des mécanismes, disons, normatifs. Ce n'est pas non plus une société dans laquelle le mécanisme de la normalisation générale et de l'exclusion du non-normalisable serait requis. On a au contraire, à l'horizon de cela, l'image ou l'idée ou le thème-programme d'une société dans laquelle il y aurait *optimisation des systèmes de différence*, dans laquelle le champ serait laissé libre aux processus oscillatoires, dans laquelle il aurait *une tolérance accordée aux individus et aux pratiques minoritaires, dans laquelle il y aurait une action non pas sur les joueurs du jeu, mais sur les règles de jeu, et enfin dans laquelle il y aurait une intervention qui ne serait pas du type de l'assujettissement interne des individus, mais une intervention de type environnemental*³⁰.

On commence alors à percevoir la rencontre non fortuite entre la gouvernementalité néolibérale et le paradigme cybernétique dont le *Big Data Stream* constitue à ce jour un des exemples les plus aboutis. Car la leçon prémonitoire de *AA Predictor* (et aussi, plus près de nous, des drones de combat et des « signatures strikes ») consiste en ceci : la technique moderne et la guerre moderne (et toute technique est une technique de guerre – ou de gouvernement des conduites) ne normalisent pas mais cherchent à composer avec, à tirer avantage des comportements individuels, ceux du pilote *et* du canonnière ; la contrainte se fait par

This self-correction is *exactly* what Wiener's machines did. Indeed, in every piece of his writing on cybernetics, from the first technical exposition of the AA predictor before Pearl Harbor up through his essays of the 1960s on science and society, Wiener put feedback in the foreground, returning again and again to the torpedo, the guided missile, and his anti-aircraft director. Moreover, even in the predictor *both* the performance *and* the rules governing performance were corrected « in the course of its functioning. » (Galison. *op. cit.*, p. 258)

30 Foucault, *op. cit.*, p. 265 ; nous soulignons.

l'environnement en jouant sur les règles qui le constituent. On est ainsi en présence d'un ajustement perpétuel dans un jeu dont « les règles ne sont pas des décisions prises par quelqu'un pour les autres³¹ ».

C'est cette conjonction essentielle entre la technique, l'économie et la gouvernementalité, celle-là même qui a guidé l'analyse du dernier Foucault, que J.-L. Nancy désigne sous le terme d'*écotechnie* dans laquelle il décèle un moment charnière marquant la fin d'une époque et l'ouverture d'une nouvelle. « Écotechnie signifie donc : l'être se révèle économiquement ; et la distinction entre la production comme concept économique et la production comme concept ontologique n'a plus cours³². » Dans ces conditions, le concept même de finalité préalable est, une fois de plus, mis à mal. Comme le note Antoine Garapon,

L'ordre spontané du marché [...] est *un ordre qui est sa propre finalité*³³.

De la même manière que la raison néolibérale ne poursuit aucune finalité, elle ne s'impose aucun objectif temporel lointain. Le monde – ne parlons plus de système – doit être *ouvert* par définition, et tout projet trop éloigné risque de nous priver de ses richesses³⁴.

Pour revenir plus spécifiquement au *Big Data Stream*, la réalité se dévoile ainsi, selon l'expression de Shoshana Zuboff, comme une nouvelle classe de marchandise : celle des comportements des êtres et des objets. Dans son travail récent, Zuboff évoque d'abord les trois « marchandises fictives » de Polanyi : la vie devenue travail, la nature devenue biens immobiliers, l'échange devenu monnaie. Elle y ajoute alors la quatrième transformation qui devient dominante au XXI^e siècle :

*Now 'reality' is subjugated to commodification and monetization and reborn as 'behavior.' Data about the behaviors of bodies, minds, and things take their place in a universal real-time dynamic index of smart objects within an infinite global domain of wired things*³⁵

31 Comme le note Foucault à propos de l'économie néolibérale, l'économie doit être un jeu, « un ensemble d'activités réglées [...] mais dans lesquelles les règles ne sont pas des décisions qui sont prises par quelqu'un pour les autres. C'est un ensemble de règles qui détermine de quelle manière chacun doit jouer un jeu dont personne, à la limite, ne connaît l'issue. » (Foucault, *op. cit.*, p. 178)

32 Frédéric Neyrat, 'Ici Le Monde', *Actuel Marx En Ligne*, 2002. Le terme de production fait ici manifestement référence au « pro-duire » heideggerien.

33 Antoine Garapon, *La Raison Du Moindre État : Le Néolibéralisme et La Justice* (Odile Jacob, 2010), p. 25.

34 Garapon, *op. cit.*, p. 39.

35 Shoshana Zuboff, 'Big Other : Surveillance Capitalism and the Prospects of an Information Civilization', *Journal of Information Technology*, 2015.

Cela veut dire exactement trois choses : que le produit du *Big Data Stream* est la réalité en tant que matière première du néolibéralisme et non en tant que « choses », « objets » ou « fonds » ; que cette réalité est le résultat de la naturalisation de comportements, eux-mêmes étant un produit dérivé des procédures complexes de traitement ; que plus ces comportements sont *modulables*, plus cette matière première est riche. On comprend alors mieux en quoi les « sociétés de contrôle » ne cherchent pas à uniformiser et à discipliner mais, au contraire, à moduler les comportements³⁶ et à favoriser les différences individuelles : comme tout « data scientist » le sait, il y a peu à tirer des données provenant d'une population uniforme.

Ainsi, dans *Big Data Stream*, les données et les algorithmes ne font qu'un au sein d'une industrie des programmes qui produisent le réel³⁷, à travers un output fondamentalement modulable et provisoire mais prédictible et opérationnel. La flexibilité à l'œuvre dans *Big Data Stream* semble correspondre à l'exigence fondamentale du paradigme cybernétique et du *machine learning* qui en est issu, celle de la recherche d'une solution sub-optimale, mais qui « marche » – c'est-à-dire qui répond à une recherche locale d'optimisation par rapport à une récompense³⁸. On quitte ici le paradigme de l'adéquation qui a guidé à la fois la science et l'interrogation philosophique de Descartes à Husserl, pour glisser vers celui de l'ajustement : ajustement des données en fonction de l'algorithme (filtrage, fréquence d'échantillonnage), ajustement de l'algorithme en fonction des données (dit « *concept drift* »), ajustement de la précision du traitement en fonction des ressources de calcul disponibles, ajustement du phénomène obtenu en fonction de l'ensemble des outputs des « capteurs virtuels ».

36 Le concept de modulation, thématiqué par Deleuze, nous semble plus à même de décrire cet état de fait de « réalité modulée » que ne le fait le concept de contrôle (y compris chez Deleuze lui-même), trop sujet à la confusion avec le contrôle dans le sens plus trivial du terme, exercé par les mécanismes dits disciplinaires et sécuritaires de Foucault.

37 Nous ne développons pas ici les différences fondamentales entre les données, les algorithmes, les programmes, le code et l'automatisation pour nous en tenir à l'équation bien connue « algorithms + data structures = programs ». Cf. à ce sujet Dourish.

38 « Optimisation / reward » correspond à l'approche dominante dans l'apprentissage par renforcement, au sein de *machine learning*. Il faut ici la mettre en parallèle avec la forme d'action environnementale chez Foucault, décrite comme l'action de modification des règles du jeu, du rapport entre les gains et les pertes possibles.

Mais il ne faut pas sous-estimer la puissance de *Big Data Stream* : il ne s'agit sans doute pas d'une simple approche « opérationnelle » mais bien de la production des savoirs, selon un paradigme nouveau dont on commence seulement à apercevoir les rouages. Et il serait illusoire de penser qu'il restera cantonné aux seuls domaines « opérationnels » qu'ils soient financiers ou publicitaires. Ces approches existent également dans les sciences de l'environnement, sans parler du fait que, au sein des Humanités Digitales, bon nombre de travaux d'analyse des interactions sociales s'appuient sur les outils *stream processing* tels que ceux de Twitter par exemple.

Le *diagramme* de *Big Data Stream* apparaît donc comme une exigence d'immédiateté, de réponse à même le flux et sa circulation autoréférentielle pour produire le réel dans le sens entendu ci-dessus. Ce diagramme perd la prétention à la capture, à la mémoire exhaustive³⁹, tant dans son aspect temporel que dans son aspect de couverture de population. Pris dans le jeu perpétuel de l'ajustement réciproque, il garde cependant, en le rejouant à sa manière, l'aspect que nous avons désigné comme absence de finalités⁴⁰.

39 Du point de vue des modèles sous-jacents, la fenêtre temporelle prise en compte par le traitement doit être d'une certaine durée optimale, mais en aucun cas infinie. Ainsi, dans les modèles Markoviens utilisés pour les traitements dits en temps réel, on remonte à n pas en arrière. La raison pour laquelle le n doit être fini n'est pas simplement celle, technique, de puissance finie de calcul, mais celle, conceptuelle, de la prédictibilité du comportement futur à partir de l'état ou d'un ensemble des états proches du présent.

40 Une approche supplémentaire et plus récente (à notre connaissance à partir de 2012), dite de *Data Lake* (lac de données), doit également être évoquée. Deux grands axes d'évolution peuvent être isolés dans l'histoire non encore écrite de *Big Data* : vers l'augmentation de la variété des données traitées et de la vitesse de traitement. La tendance à la variété, en un certain sens, reflète et accentue l'exigence de l'exhaustivité mémorielle dont nous avons parlé. En effet, pour augmenter la performance analytique, les données non-structurées ou poly-structurées sont capturées et stockées directement sous le format natif dans les *Data Lakes*. : « Data lakes operate [...] with a deluge of information being diverted into a large repository where it can be held for a long period of time until it is needed. The lakes are constantly fed by new water flowing into the lake, and in fact, they are dependent on the constant flow of water to keep the environment vibrant otherwise the lake could stagnate. » (Chimsky, *op. cit.*). Ce souci d'assurer le flux constant, de ne pas laisser les données stagner et devenir obsolètes reflète aussi, à l'état naissant, l'exigence de la rapidité de traitement. Mais *Data Lake* offre une approche encore trop statique et diffère de *Data Stream* sur des points essentiels : « Analysis using data streams is a fundamentally different approach than data lakes. Rather than diverting the flow to store and then analyze, with streams, analysis occurs as the information is flowing in real- or near-real time. » (Chimsky, *op. cit.*) – Ainsi, *Data Lake* présente une variante

LES DEVENIRS DE L'ARCHIVE

La qualification du phénomène Big Data comme exhaustivité mémorielle *et* immédiateté amène, à notre sens, à poser la question du pouvoir à de nouveaux frais. Chez Foucault, nous l'avons déjà évoqué, le ressort interne des transformations historiques ne réside pas tant dans un diagramme particulier que dans la tension entre des diagrammes simultanés, l'exemple le plus connu étant la tension entre les mécanismes disciplinaires et de sécurité⁴¹. De même, dans Big Data, c'est finalement le contraste entre le *Warehouse* et le *Stream* qui fait apparaître la tension. Ce nœud de tension n'est cependant pas tant dans la nouveauté technique ou scientifique en tant que telles – après tout un changement de paradigme peut être nécessaire ou bienvenu – mais dans la redistribution des rapports entre l'archive et le diagramme que nous allons maintenant esquisser.

Nous l'avons vu, le deuxième diagramme comporte des caractéristiques particulières. Le caractère autoréférentiel de la circulation des données qu'il donne à voir, la décentralisation de l'analyse vers les capteurs à la place de l'instance centralisée, la modification permanente des règles d'analyse et de production de l'environnement, le parallélisme sans réserve avec les forces à l'œuvre dans la gouvernementalité néolibérale : l'image qui se dessine ainsi fait penser à l'espace lisse de la circulation des forces, réifiées ici à travers les flux de données qui ne demandent qu'à être analysées. Avec le *Big Data Stream*, on pourrait alors être amené à envisager quelque chose comme un *pur diagramme*, un pur rapport de forces, qui devient visible et opérant en effaçant ou en lissant le strié de l'archive.

En un certain sens, il en est en effet ainsi, et il se pourrait bien que, tendanciellement, l'archive soit amenée à disparaître.

Quel est le rôle de l'archive qui semble ici condamnée à disparaître ? Chez Deleuze, comme nous l'avons évoqué, le savoir est un vis-à-vis du pouvoir. En revenant à leur présupposition réciproque, rappelons que si

intermédiaire entre le *Warehouse* et le *Stream* (*modulo* le problème de la structuration des données qui dépasse le cadre de cet article).

41 Cf. note 5 ci-dessus.

la constitution des savoirs présuppose et actualise les rapports de pouvoir, le pouvoir ne se donne à voir qu'à travers les savoirs qu'il sous-tend. En évoquant la société « disciplinaire », Deleuze utilise la métaphore de la taupe, image du pouvoir qui ne voit pas mais fait voir : la taupe ne se reconnaît qu'à travers le réseau de ses galeries, des savoirs stratifiés que sa circulation produit et rend visibles. Mais si la taupe se reconnaît à travers ce réseau, c'est aussi lui qui la rend visible. L'archive est alors ce qui rend possible un discours, une énonciation sur les rapports de pouvoir à l'œuvre, ce qui les met au jour. Qui plus est, et c'est en cela peut-être que l'on peut prendre une certaine distance avec Deleuze, on ne modifie la circulation de la taupe qu'en modifiant les strates, l'archive elle-même. C'est l'archive qui permet un recul, une prise de position critique et une prise sur le pouvoir. En d'autres mots encore, si le diagramme est une représentation qui rend visibles les forces et les rapports de pouvoirs, les conditions de possibilité d'une *prise* sur les forces qui y circulent résident dans l'archive, et ceci en raison même de leur présupposition réciproque.

En poursuivant l'hypothèse du pur diagramme dans *Big Data Stream*, ces conditions de possibilité sont-elles encore réunies ? La tension entre le *Stream* et le *Warehouse* ne fait-elle pas apparaître que, paradoxalement, c'est du côté de ce dernier que la place existe peut-être encore pour une constitution de l'archive, un temps de suspension, un recul critique nécessaire, une prise sur le pouvoir ? La nouveauté, dont cette tension est le vecteur, consisterait alors dans la dissolution progressive de l'archive au profit d'un pur rapport de forces immédiatement visible et opérant.

Mais, en un autre sens, en revenant plus près de la conception deleuzienne du rôle structurel de la présupposition réciproque, le *pur diagramme* sans l'archive n'est pas concevable. Et il ne faut pas s'y tromper, c'est bien du savoir (certes sur un mode nouveau), du strié et donc de l'archive qui est produit par le *Big Data Stream*. Seulement, cette archive est dans un rapport nouveau au diagramme, un rapport qui n'est plus de l'ordre d'un vis-à-vis mais d'une simple soumission. Nous sommes alors dans une situation nouvelle dans laquelle le diagramme prend le pas sur l'archive : non dans le sens où cette dernière disparaît, mais dans le sens où le pouvoir n'en dépend plus, où il peut désormais « se conserver et se reproduire⁴² » sans être actualisé par les savoirs. Ainsi,

42 Cf. la note 1 ci-dessus.

le diagramme gagne ici en invisibilité, en exposant l'indéchiffrabilité de la circulation des données, l'autonomisation par rapport à l'archive, l'auto-actualisation indéfinie⁴³.

Pour conclure, la question du pouvoir pose de toute évidence celle du contre-pouvoir, de la résistance. En résumant ici de manière excessivement schématique, il nous semble que Deleuze et Guattari actent le passage de la résistance centralisée, arme contre les mécanismes disciplinaires eux-mêmes centralisés, aux « luttes transversales » ou « moléculaires ». Ils inaugurent ainsi toute une lignée de pensée et de penseurs qui va élaborer deux variantes de ces nouvelles formes de résistance. Pour employer ici les termes de Deleuze lui-même, celles-ci peuvent prendre la forme des micropolitiques qui permettraient d'agir à la marge en échappant aux circuits principaux de pouvoir ou en les déstabilisant, ou bien la forme de « vacuoles de non-communication » qui permettraient des temps de suspension et de recul nécessaires. La question se pose désormais de savoir si l'une ou l'autre de ces variantes possède une « force » suffisante au sein du diagramme de *Big Data Stream*, où tous les circuits semblent investis par le pouvoir, où la suspension devient de plus en plus difficile et où le recul critique est remis en cause par la relation nouvelle avec l'archive. En d'autres termes encore, peut-on penser des résistances en dehors du paradigme de l'archive ?

Armen KHATCHATOUROV
Télécom École de Management
IMT – Institut Mines-Télécom

⁴³ Comme on le sait, Deleuze conclut son *Post-scriptum* en utilisant la métaphore du serpent et de la circulation indéfinie, et en opposant le serpent, image des sociétés de contrôle, à la taupe des disciplines : « Les anneaux d'un serpent sont encore plus compliqués que les trous d'une taupinière ». Gilles Deleuze, 'Post-Scriptum Sur Les Sociétés de Contrôle', in *Pourparlers 1972-1990* (Les éditions de Minuit, 1990).