



INSTITUT  
DIDEROT

Les Carnets des Dialogues du Matin

---

KENNETH CUKIER - DOMINIQUE LEGLU

# L'avenir des Big-Data

---

Les Carnets des Dialogues du Matin

---

KENNETH CUKIER - DOMINIQUE LEGLU

# L'avenir des Big-Data

---

PRINTEMPS 2015

---

# Sommaire

L'avenir des Big-Data p. 5  
Kenneth Cukier

Débat p. 13  
Kenneth Cukier et Dominique Leglu

Discussion p. 19

Postface p. 31  
Dominique Lecourt

Les publications  
de l'Institut Diderot p. 39

---

# L'avenir des Big-Data

---

Les *big-data* constituent une véritable révolution. La croissance exponentielle des données détermine en effet une nouvelle façon de les employer. Celles-ci sont désormais susceptibles d'une multitude de réutilisations sans rapport avec leur collecte initiale. La révolution des *big-data* représente ainsi bien plus qu'un changement quantitatif : en modifiant notre façon de traiter les données, elle bouleverse notre représentation du savoir, de l'expertise, de l'homme et de catégories juridiques comme celles de vie privée ou de responsabilité.

Pour comprendre en quoi consiste cette révolution, posons les trois questions suivantes :

- comment faire pour mesurer les précipitations si on ne dispose pas d'une machine spécifiquement conçue à cet effet ?
- comment établir une contre-indication pour un médicament dans le cas où il n'est pas possible de mener une étude médicale ?
- comment savoir à quel moment acheter un billet d'avion sur internet ?

---

Dans les trois cas, les *big-data* ont permis de résoudre la question. Non seulement grâce au vivier de données que les progrès technologiques ont permis de constituer, mais aussi, et surtout, parce qu'on a imaginé des façons nouvelles et astucieuses d'utiliser les données existantes.

Prenons la troisième question : c'est celle que s'est posée Oren Etzioni, le fondateur de *Farecast*. Ce professeur d'informatique à l'Université de Washington s'est rendu compte à l'occasion d'un vol Seattle-Los Angeles qu'il avait payé son billet plus cher que son voisin. Etzioni avait pourtant acheté sa place très longtemps à l'avance, bien avant son voisin et d'autres passagers auxquels il a posé la question. Il ne suffit donc pas d'acheter son billet longtemps à l'avance pour l'avoir au moindre prix – la formule est plus complexe, et fait intervenir une myriade de facteurs que les compagnies aériennes gardent pour elles. Etzioni a alors mis ses compétences d'informaticien à l'œuvre afin d'élaborer, à partir des données fournies par des sites d'achats de billets, un algorithme permettant de prédire le meilleur moment pour acheter un billet sur un vol donné. Cela a donné *Farecast*.

En quoi cette histoire illustre-t-elle la révolution introduite par les *big-data* ? Premièrement, parce qu'elle n'a été possible que grâce à la révolution numérique. 200 milliards de données concernant les prix de presque tous les vols de l'aviation civile américaine, sur toutes les routes, pour tous les sièges et toutes les villes, ont été traitées. Le projet *Farecast* repose sur une quantité

---

gigantesque de données qu'il était impossible ni d'avoir ni de maîtriser ne serait-ce qu'il y a dix ans. Ensuite, Etzioni a réutilisé de façon astucieuse des données qui avaient été collectées à l'origine pour une autre fin, établir une transaction commerciale et constituer des dossiers clients. Etzioni a ainsi élaboré une méthodologie spécifique, qui peut-être réutilisée dans de nombreux autres domaines présentant des similarités avec les billets d'avion, comme les chambres d'hôtel, ou les billets de concert. Enfin, Etzioni s'est contenté de repérer les corrélations permettant de prévoir les variations dans les prix de billets, sans chercher à connaître la raison de ces variations. *Farecast* se contente du fait de la variation, sans chercher à percer le « pourquoi ? ».

La révolution des *big-data* suppose donc, bien entendu, une capacité inédite à capter, conserver et traiter les données. Mais ce qui frappe dans le cas de *Farecast* est le fait que ces données ont été réutilisées à des fins nouvelles. Car quand les systèmes électroniques de vente de billets ont été mis en place à partir des années 60, les compagnies aériennes n'ont pas imaginé que leurs données puissent être retournées contre elles, c'est-à-dire contre leur *business model*. Il y a donc au cœur des *big-data* un certain état d'esprit, un « esprit *big-data* », qui consiste à voir dans les données une valeur latente qu'on n'avait pas imaginée au moment de les collecter. Il est important de comprendre que les données sont devenues une véritable ressource. Les économistes classiques n'ont jamais imaginé qu'en plus de la terre, du travail et du capital les données puissent

---

constituer un facteur de production, cela pour la simple raison que collecter les informations était difficile et demandait beaucoup trop de temps. Mais ce n'est plus le cas aujourd'hui : il est maintenant infiniment plus facile de collecter et de manipuler les données, et il est possible pour cette raison d'y trouver des richesses insoupçonnées auparavant.

La révolution des *big-data* ne se réduit donc pas à une simple explosion quantitative. Il est indéniable que le passage à l'ère numérique a eu pour conséquence une augmentation exponentielle et vertigineuse des données dont nous disposons, ainsi que de nos capacités de stockage. La quantité d'informations dont nous disposons ferait plus que doubler tous les trois ans, et serait en 2013 d'environ 1 200 exaotets, ou 1200 milliards de milliards d'octets, dont moins de 2 % sous forme non numérique. Dès l'origine, cependant, le terme renvoie aux implications de cette croissance, plutôt qu'à une simple description quantitative. L'expression en effet est apparue il y a une dizaine d'années pour décrire le fait que la quantité de données avait à ce point explosé que l'infrastructure technologique, *hardware* ou *software* ne suffisait plus pour les traiter. Il devenait donc nécessaire de modifier cette infrastructure. L'expression « *big-data* » renvoie ainsi, au début, au problème technologique posé par la quantité inédite de données à traiter. Mais c'est aussi à cette époque que l'on s'est aperçu que l'apprentissage automatique des machines, qui jusqu'ici ne fonctionnait pas très bien, devenait beaucoup plus efficace avec une injection

---

massive de données. Des progrès considérables ont ainsi été faits dans la conception de programmes de jeu d'échecs ou de traduction. Ces progrès ne sont toutefois pas le résultat d'une avancée dans la conception des algorithmes ou de l'amélioration des micro-processeurs. Ce qui a tout changé, c'est l'ajout massif de données. Il en va de même quand il s'agit de mettre au point des voitures à conduite automatique : des progrès ont été faits grâce à l'utilisation d'un grand nombre de données plutôt qu'à une modification de la conception des logiciels. L'expression « *big-data* » a ainsi commencé à être employée pour renvoyer à une seconde idée : *more is not more, more is different* – le changement quantitatif aboutit à une évolution qualitative.

Dans le livre que j'ai publié avec Viktor Mayer-Schönberger <sup>1</sup>, nous résumons donc la révolution des *big-data* en trois mots : *more, messy, correlation*. Ce sont les trois piliers du *big-data*.

*More* : nous disposons d'infiniment plus de données qu'auparavant. Nous n'avons donc plus besoin de nous contenter d'échantillons, car nous pouvons avoir toutes les données relatives à un phénomène particulier. Cela nous permet, tout d'abord, d'éviter les difficultés techniques posées par la construction de l'échantillon. Mais le fait de disposer d'une quantité gigantesque d'informations autour d'un phénomène,

---

1. Viktor Mayer-Schönberger & Kenneth Cukier, *Big-Data - La révolution des données est en marche*, trad. franç. H. Dhifallah, Paris, Robert Laffont, 2014.



---

voire de *toutes* les informations, permet surtout deux choses. Premièrement, de ne pas relancer les processus de collecte de données si l'on se rend compte que l'échantillon ne couvre pas les données dont nous avons besoin, par exemple parce que nous venons d'imaginer une nouvelle façon de les employer. Ensuite, d'améliorer considérablement la granularité de l'analyse : les *big-data* permettent de repérer et de suivre des sous-groupes, de taille très réduite, de façon bien plus efficace que l'échantillonnage.

*Mesyy* : il s'agit d'un terme technique, que l'on peut traduire par « en désordre », ou « imprécis ». On distingue ainsi les *messy data* des *clean data*, des données propres, bien ordonnées et bien construites. Dans le passé, parce que les quantités de données étaient faibles, il était nécessaire que celles-ci soient les plus propres possibles. Sinon leur extrapolation peut s'avérer catastrophique. Il y a en effet le risque, si l'on n'a que peu de données, d'amplifier les erreurs, avec des conséquences graves. Prenons l'exemple de la température dans un vignoble. Si l'on ne dispose que d'un seul thermomètre pour des centaines de pieds de vigne, ce thermomètre doit fonctionner en permanence et être complètement fiable. Un dérèglement, ou un arrêt temporaire, peut entièrement fausser la mesure de la température, et avoir des conséquences dramatiques pour les vignerons. En revanche, si l'on dispose d'une centaine de capteurs, un par pied de vigne, l'ensemble des données pourra être relativement imparfait, car même si tel ou tel capteur peut dysfonctionner, les données seront au

---

total fiables et même plus riches d'enseignement. On voit là que l'imprécision des données est très largement compensée par leur nombre. Quand on dispose de quasiment toutes les données autour d'un phénomène, on peut tolérer des données imparfaites. Les données « propres » restent préférables, bien entendu, mais on peut tolérer que celles-ci soient en partie imprécises.

*Correlation* : Pendant longtemps, le but pour un savant a été d'expliquer le monde, de trouver la causalité cachée derrière les phénomènes. Les *big-data* nous permettent de repérer des corrélations, et d'agir à partir d'elles, sans que nous ayons à nous préoccuper de la raison de ces corrélations. Les *big-data* nous enseignent l'humilité : le monde est complexe, il est difficile de trouver une causalité indubitable, qui ne sera pas remise en cause dans l'avenir par l'analyse de nouvelles données, ou d'un plus grand nombre d'entre elles. Les corrélations se contentent de repérer un « quoi » : quelles variables sont corrélées, et de quelle façon, sans chercher derrière le « pourquoi ». Cela suffit amplement pour nous faire voir des phénomènes intéressants, et pour nous aider à agir. Ainsi, les spécialistes en données, en statistiques, ou en traitement automatisé de l'information remplaceront probablement en grande partie les experts d'un champ donné. Pour ne donner qu'un exemple : lors d'une récente étude conduite à Stanford, une machine s'est révélée capable d'identifier sur un prélèvement des indices de la présence d'un cancer auquel les humains n'avaient pas pensé. Les experts ne disparaîtront pas, mais les *big-data* entraîneront certainement une redéfinition des tâches.

---

Les *big-data* et les nouvelles technologies élimineront ainsi des métiers et redéfiniront le monde du travail, comme l'ont fait la chaîne et l'automatisation au siècle dernier. Ce n'est pas la moindre des conséquences de cette révolution.



---

## DÉBAT AVEC DOMINIQUE LEGLU <sup>2</sup>

**Dominique Leglu** : *Je voudrais revenir sur l'expression « more is different » : en quoi s'applique-t-elle aux questions que vous avez données au début ? Par exemple, que faire quand on ne dispose pas de système météo permettant de savoir où il pleut et avec quelle force ? Peut-on arriver à le savoir, grâce à des données imprévues ?*

**Kenneth Cukier** : Le principe est le suivant : les antennes-relais ne communiquent pas seulement avec les téléphones, mais aussi entre elles. Or une chose qu'elles se communiquent est leur consommation d'énergie, et la pluie perturbe cette consommation d'une façon que l'on sait mesurer. En observant ces détériorations, on peut donc savoir s'il pleut, et en quelle quantité. Bien entendu, ce procédé n'est pas nécessaire en Europe ou aux États-Unis, car on dispose de stations météo, et on peut les mettre là où on souhaite mesurer les précipitations. Mais dans d'autres régions, en Afrique notamment, il y a des problèmes de financement, certains endroits sont difficiles d'accès, loin des villes, tout en étant inondables. Utiliser les *big-data* fournies par les antennes-relais peut alors devenir très précieux.

**Dominique Leglu** : *Les opérateurs de téléphonie mobile peuvent donc nous donner des informations sur la météo...*

---

2. Dominique Leglu est directrice de *Sciences et Avenir*.

---

**Kenneth Cukier** : Tout à fait – on n’en est pour l’instant qu’au stade expérimental, mais cette technologie pourra dans quelques années être mise en place.

**Dominique Leglu** : *Vous avez aussi mentionné le cas des industries pharmaceutiques. Avant de mettre sur le marché un nouveau médicament, elles doivent mener pendant des années des essais cliniques pour étudier ses effets en combinaison avec d’autres substances. Mais les entreprises pharmaceutiques ne peuvent étudier l’ensemble de ces interactions. Pourront-elles alors se tourner vers d’autres acteurs qui disposeraient grâce aux big-data d’informations supplémentaires ?*

**Kenneth Cukier** : Oui, et il est assez simple d’imaginer comment. Un article est paru sur la question en 2013.<sup>3</sup> La méthode consiste à étudier les millions de requêtes internet portant sur un médicament, sur un second médicament, et enfin sur les deux médicaments ensemble. Il apparaît alors que les requêtes associant les deux termes sont corrélées à une augmentation des requêtes sur l’hyperglycémie. L’usage de ces deux médicaments semble donc présenter un risque accru d’hyperglycémie. Ce qui est intéressant, c’est que cette contre-indication était inconnue des entreprises pharmaceutiques ou des autorités médicales américaines. Des sociétés disposant de ce type de données pourraient donc les vendre à ceux qui seraient intéressés par ce type de corrélation.

---

3. White RW & al. “Web-scale pharmacovigilance : listening to signals from the crowd”, *J Am Med Inform Assoc.*, mai 2013, vol. 20, n°3, pp. 404-408.

---

Par exemple, des sociétés d'assurance. On voit ainsi comment il est possible de repérer grâce aux *big-data* des phénomènes importants, pour peu que l'on en fasse un usage astucieux.

**Dominique Leglu :** *C'est un très bon exemple de corrélation : les gens prennent tel médicament sans savoir, parce qu'il n'y a pas eu tous les essais nécessaires, qu'il est contre-indiqué en association avec un autre médicament, et font par la suite des recherches sur internet parce que quelque chose ne va pas. La grande nouveauté est que ceux qui disposent des données relatives aux requêtes sur les moteurs de recherche vont pouvoir alors donner cette information. Mais il ne s'agit que d'une corrélation, on ne sait pas à quel phénomène biologique précis correspond cette interaction entre les deux médicaments.*

**Kenneth Cukier :** Il faut toutefois bien comprendre qu'il ne s'agit là pas d'une preuve irréfutable. La corrélation n'indique pas avec certitude la présence d'un problème, mais nous indique où mener des recherches. Le *Big-Data* ouvre la recherche, au lieu de la clore.

**Dominique Leglu :** *Vous commencez votre livre par une anecdote bien connue, qui a révélé le potentiel des big-data au grand public, à savoir le fait que les requêtes sur Google ont pu permettre de suivre en temps réel, ce que ne permettent pas les méthodes classiques, la progression de la grippe aviaire aux États-Unis en 2009. Or ces requêtes n'étaient pas nécessairement très précises, et les données étaient donc messy...*

**Kenneth Cukier :** Un article vient cependant d'être

---

publié dans Science <sup>4</sup>, après la publication de notre livre, qui critique assez fortement *Google Flu Trends* et les *big-data*. En forçant le trait, cet article affirme que les *big-data* ne marchent pas. Et de fait *Google Flu Trends* a pu fonctionner un moment, mais semble se montrer bien moins efficace depuis deux ou trois ans. Je crois qu'il faut être prudent, dans un sens comme dans l'autre. La presse brûle maintenant ce qu'elle a adoré il y a juste quelques années. Je dirais pour ma part que quand *Google Flu Trends* surestime les épidémies de grippe ces dernières années, ce n'est pas très différent de ce qui arrive quand la météo annonce de la pluie et qu'il ne pleut pas. Il s'agit de probabilités, pas de certitudes.

**Dominique Leglu** : *Je voudrais revenir sur la valeur latente des données : est-ce que la réutilisation des données est infinie, et va-t-il falloir tout garder, tout le temps ?*

**Kenneth Cukier** : Les données sont désormais susceptibles d'une myriade de réutilisations, impossibles à prévoir au moment de leur collection. Ces réutilisations ne sont peut-être pas infinies, mais cela représente quand même quelque chose de fondamentalement nouveau. Dans le livre, nous prenons l'exemple de Walmart, qui collecte des données pour suivre ses ventes mais qui surtout les croise avec d'autres bases de données afin de trouver des corrélations susceptibles d'une exploitation commerciale. En croisant les données météo avec ses

---

4. David Lazer & al., "The Parable of Google Flu : Traps in Big-Data Analysis", Science, mars 2014, vol. 343, n°6176, pp. 1203-1205.

---

chiffres de ventes, Walmart s'est ainsi aperçu qu'avant une tempête les clients avaient tendance à acheter des *Pop-Tarts*. Les magasins mettent donc des *Pop-Tarts* près des caisses avant une tempête, et les ventes ont augmenté. Le plus intéressant, dans l'affaire, est qu'il ne s'agit pas là de piles électriques ou de lampes torches, c'est-à-dire de quelque chose qu'on aurait pu deviner *a priori*. Ce n'est là qu'une des multiples façons dont on peut réutiliser les données dont on dispose, pour peu que l'on ouvre les yeux.

**Dominique Leglu** : *Vous n'avez pas évoqué les données fournies par nos cerveaux... Il y a là bien évidemment un risque d'atteinte à la vie privée. Or dans votre livre, vous insistez sur le fait que selon vous le risque n'est pas tant celui d'un Big Brother, d'un pouvoir totalitaire, que d'un scénario à la Minority Report, où toutes ces données seraient utilisées pour prédire le comportement de chacun, et aboutiraient à des arrestations préventives avant qu'un crime ne soit commis.*

**Kenneth Cukier** : Bien entendu, le risque d'un *Big Brother* est à prendre au sérieux, mais ce n'est pas un problème propre aux *big-data*, même si ceux-ci le rendent probablement plus menaçant. Nous insistons dans le livre sur un autre danger : les algorithmes permettront de faire des prédictions sur la dangerosité des personnes, ou leur probabilité de commettre un acte particulier, à partir des *big-data* et des corrélations que l'on peut y trouver. Un individu pourrait être l'objet de mesures de contrôle, stigmatisantes de surcroît, et même être pénalisé avant même d'avoir agi dans le



---

mode réel. Et dans le même temps il sera difficile de résister devant les prévisions apportées par les *big-data*. Que faire si un algorithme affirme qu'une personne a une probabilité de 95 % de commettre un vol dans les 6 prochains mois ? Si on l'arrête et qu'on le met en prison, on viole sa capacité à faire des choix, et le principe fondamental selon lequel nous ne devons être sanctionnés que pour ce que nous avons effectivement fait. Mais refuser d'utiliser ces résultats c'est refuser la science, s'exposer à des dommages qu'on aurait pu éviter, et dont les victimes pourraient nous faire grief. Notre système juridique n'a pour l'instant pas de réponse à ce sujet – c'est tout l'intérêt des *big-data* : ils nous obligent à repenser notre droit, dans le domaine de la responsabilité comme dans celui – nous en parlons dans le livre – du respect de la vie privée, les moyens traditionnels comme l'anonymisation ou la notification-consentement étant désormais très peu efficaces.



---

## DISCUSSION AVEC LA SALLE

**Dominique Lecourt** : *Votre exposé ouvre à mon sens deux grandes pistes de réflexion. La première est épistémologique : les big-data iraient jusqu'à remettre en question notre conception de la connaissance. Il faudrait abandonner non seulement la question du « pourquoi », comme vous le dites explicitement, mais aussi la question du « comment » qu'on a tenté au XIX<sup>e</sup> siècle de substituer à la conception classique de la science centrée autour de la causalité. La question à poser serait désormais celle d'un simple « quoi » : quelles corrélations apparaissent à l'analyse des données ? Il y a là un chantier immense pour les philosophes des sciences. La seconde grande piste de réflexion que je vois est la suivante : les big-data nous demandent de réviser notre conception non seulement de la science, mais aussi de l'homme. Celui-ci apparaît moins singulier qu'il se flatte d'être en tant qu'individu, et il est prévisible. Dans ce cas, comment éviter, comme cela a été dit, un scénario à la Minority Report ? L'impact éthique et politique des big-data est immédiat, et nous n'avons pour l'instant pas de réponse aux problèmes qu'ils posent.*

**Kenneth Cukier** : Ce ne serait pas la première fois que notre conception du savoir ou de notre place dans le monde serait bouleversée... L'ironie avec les *big-data* est que contrairement à ce qu'on pourrait croire, ils peuvent nous inciter à une certaine modestie face à la vision moderne de la science qui s'imagine que tout peut être connu et expliqué scientifiquement grâce aux mathématiques, ce qui donne des pseudo-sciences comme la phrénologie. Les *big-data* nous montrent que si on étudie de grands ensemble de données, il apparaît

---

que nous en savons moins que nous pensons.

**Dominique Leglu :** *Ce que vous dites me refait penser à un exemple assez extraordinaire que vous donnez dans le livre, celui des prématurés...*

**Kenneth Cukier :** ... une équipe au Canada a en effet cherché à collecter le maximum d'informations possibles concernant les bébés prématurés, afin de mieux prédire l'apparition d'une infection. Or ce qu'ils ont trouvé, c'est que les meilleurs prédicteurs d'une infection ne sont pas ceux que l'on aurait cru au départ. Auparavant, il s'agissait de voir à intervalles réguliers si les signes vitaux étaient stables, l'idée étant qu'une détérioration de ces signes vitaux précède l'infection. Il arrivait ainsi que le médecin regarde l'état du prématuré en fin de journée, constate qu'il est stable, et rentre chez lui pour apprendre dans la nuit que le bébé est décédé. Or ce qu'ont montré Carolyn McGregor et son équipe de l'Institut de technologie de l'Ontario, c'est qu'il est possible de détecter 24h à l'avance et avec une plus grande fiabilité une infection chez le prématuré si l'on repère dans le flux d'informations vitales qu'il transmet d'infimes changements, impossible à voir l'œil nu mais décelables par ordinateur.

**Dominique Leglu :** *Il s'agit ainsi de suivre en permanence le système vital du prématuré, en mesurant de façon continue tous les paramètres envisageables, comme le battement du cœur, le rythme respiratoire, la tension artérielle ou le taux d'oxygène dans le sang. Le bébé est ainsi constamment contrôlé au moyen de dizaines et de dizaines de données captées en permanence.*

---

**Kenneth Cukier** : De plus, ces flux de données peuvent être conservés et réemployés, alors qu'auparavant nous étions contraints de suivre certaines données seulement, par exemple la température, et à s'en débarrasser après traitement. Or il est possible maintenant de suivre une multitude de données, et surtout de les conserver pour les réanalyser et les réutiliser d'une nouvelle façon. Dans ce domaine, les pays comme la France qui ont un système médical centralisé ont un avantage certain. Les États-Unis en revanche ont beau avoir l'esprit *big-data*, les compétences et les moyens d'investir, ils partent néanmoins un handicap...

**Dominique Leglu** : *Une précision : la Sécurité Sociale conserve pour l'instant ses données, elle refuse d'en faire des open data...*

**Philippe Even** : *Mon sentiment est que dans l'expression « big-data », ce n'est pas le « big » qui pose problème, car il s'agit d'une difficulté technique que nous arriverons à résoudre, mais ce sont les « data ». La revue The Lancet vient de consacrer, ce qui est tout à fait inhabituel, 80 pages à 5 études signées par des grands noms comme John Ioannidis à Stanford ou Iain Chalmers à Oxford<sup>5</sup>. Ces articles analysent la situation actuelle de la recherche biologique et médicale. La conclusion est que 80 % des informations publiées ne sont pas reproductibles ou même fausses pour deux grandes raisons. Lorsqu'il s'agit de la recherche privée, celles de l'industrie pharmaceutique par exemple, c'est parce qu'elles sont délibérément falsifiées ou*

---

5. *The Lancet*, vol. 383, n°9912 et n° 9913, 11 et 18 janvier 2014.

---

*gardées secrètes. Pour ce qui est de la recherche publique, les données ne sont pas falsifiées, ou très rarement, mais elles sont intrinsèquement invalides pour toutes une série de raisons qui sont très bien analysées dans ces articles, qui pointent en particulier la responsabilité des administrations de la recherche qui accordent les financements, et le fait que les chercheurs sont obligés de passer les trois quarts de leur existence à trouver des fonds. Il me semble donc que l'un des problèmes majeurs pour les big-data est celui de la fiabilité des données et de leur reproductibilité. Ceci pose notamment la question de la publicité de ces data. Or l'industrie pharmaceutique garde pour elle l'ensemble des informations qu'elle recueille et elle continue de refuser d'ouvrir ses dossiers malgré les pressions dans tous les pays du monde.*

**Kenneth Cukier** : On pourrait croire que les *big-data* offrent des solutions à ce problème de la fiabilité des données et de la fraude. Je ne le pense pas. Il s'agit d'un problème humain, et cela ne changera pas avec les *big-data*. Il n'y a pas d'obligation à partager les données dans un monde de *small data*, et il n'y a pas de raison que cet impératif s'impose de lui-même avec les *big-data*. Le *big-data* peut changer des choses, mais ne nous sauvera pas si quelqu'un fait une fraude. Les *big-data* posent même un danger supplémentaire, celui des corrélations fallacieuses, dues simplement au hasard, étant donné la masse de données et de variables.

**Bernard Granger** : *Les données de la Sécurité Sociale en France, qui sont d'ailleurs corrélées avec celles des séjours hospitaliers, sont un instrument exceptionnel, c'est une des bases de données les plus importantes au monde, mais malheureusement*

---

*elle est très largement sous-utilisée. Quelques études ont été faites, par exemple sur le traitement de l'adénome de la prostate, qui ont montré que certaines procédures chirurgicales donnaient beaucoup plus de complications que d'autres, ou que les médecins généralistes étaient aussi performants pour le suivi médical des patients que les médecins spécialistes. Je suis convaincu que l'utilisation de ces données pourrait être un outil majeur pour le pilotage des politiques de santé publique, l'organisation de soins, la mise en place de parcours de soins optimaux. J'ai tenté de développer ces arguments devant les parlementaires, les technocrates de la santé, la ministre, et à chaque fois j'ai eu l'impression d'être un extraterrestre... Comment expliquer cette inertie des décideurs relativement à ces outils de décisions qui sont beaucoup plus puissants que ceux utilisés actuellement ?*

**Kenneth Cukier :** Les hommes politiques et les administrateurs ne voient pas toutes les possibilités offertes par les *big-data*, or je pense que dans les prochaines années leur emploi sera aussi banal qu'internet aujourd'hui. Il faudra probablement commencer par parler en termes comptables : les responsables seront tout ouïe si on leur montre que les *big-data* permettent de diminuer les dépenses de santé. Il faudra tenir deux discours : dire d'abord qu'un usage efficace des données permet de diminuer les coûts, pour montrer ensuite qu'il permet de servir l'intérêt général. La France a les moyens de devenir un centre d'excellence dans le domaine des données et de la santé, à la façon dont la Silicon Valley est au centre de l'Internet. Encore faudra-t-il faire preuve de vision, et ne pas reproduire l'erreur du Minitel...

---

**Gilles Benoît** : *J'aurais voulu avoir votre sentiment sur l'utilisation des big-data dans la finance, en particulier chez les hedge funds qui sont, je pense, particulièrement friands des big-data et croient pouvoir s'en servir pour maximiser leurs profits.*

**Kenneth Cukier** : Les fonds d'investissement ont en effet recours à ce que l'on nomme le « *trading algorithmique* ». Ce terme peut renvoyer à deux choses : le *trading* à haute fréquence, ou le fait d'utiliser de manière astucieuse beaucoup de données. Le premier pose des problèmes spécifiques, notamment de délits d'initié. Mais de façon générale le *trading* algorithmique pose deux problèmes : d'une part, ce type de trading automatisé a pour conséquence de faire exploser les régulations et la théorie de la fixation des prix par des agents rationnels ; d'autre part, les *big-data* permettent d'identifier des corrélations complexes et d'élaborer des modèles super-sophistiqués, qui sont en fait plutôt idiots et qui constituent l'une des raisons de la crise économique de ces dernières années.

**Valérie Planchet** : *Les corrélations dont vous avez parlé posent le problème de la fiabilité des données. Car celles-ci peuvent être manipulées. Si on reprend l'exemple des crânes, on sait aussi que pour montrer la supériorité des blancs sur les noirs on prenait pour les seconds des crânes de femmes qui sont généralement plus petits puisque les femmes ont une taille moins élevée. On voit bien que les corrélations qu'on trouve dépendent aussi de ce qu'on essaye de montrer. De plus, au-delà de la question de la manipulation des données, leur fiabilité aussi est discutable. Car à partir du moment où les données deviennent*

---

*publiques et transparentes, le fait de savoir ce que les autres pensent modifie nos propres choix. On peut prendre l'exemple de l'achat de livres où le fait de savoir ce que les autres ont préféré conduit les gens à leur tour à ne lire que les livres qui sont déjà connus, et qui ne sont pas forcément ceux qu'ils auraient choisis à l'origine. Au lieu de simplement repérer des corrélations statistiques, le Big-Data ne risque-t-il pas plutôt d'aboutir à des prédictions auto-réalisatrices ?*

**Kenneth Cukier** : Il s'agit là en effet d'un danger important. Je voudrais préciser que notre livre n'est certainement pas une apologie des *big-data* : nous essayons dans notre livre de présenter les périls qui y sont associés, et dont il faut bien avoir conscience si l'on veut mettre en place des moyens d'y remédier.

**Dominique Leglu** : *Vous signalez aussi dans le livre que les big-data entraîneront nécessairement la création de nouveaux métiers. La question de la gouvernance des big-data deviendra essentielle : que ce soit dans les entreprises, ou chez les régulateurs une fois qu'ils auront compris que ça existe, il faudra trouver des moyens d'éviter les fraudes que vous décrivez.*

**Claude Got** : *Ce qui est fascinant mais aussi inquiétant dans les big-data c'est la possibilité qu'elles offrent de repérer de nombreuses corrélations sans donner d'explication de type causal. En croisant les données fournies par la Sécurité sociale et les procès-verbaux enregistrant les accidents de la route, c'est-à-dire une masse de documents couvrant des centaines de milliers d'accidents sur des dizaines d'années, une équipe de Bordeaux a étudié les corrélations entre accidents et médicaments psychoactifs.*



---

*Toute une série de corrélations apparaissent, mais ensuite ? Une fois qu'on a montré qu'un malade psychiatrique ou qui reçoit des médicaments psychoactifs a un niveau d'accidentalité accru, il n'y a plus grand-chose à dire, si on en reste à la simple identification d'une corrélation. Il me semble que même si on dispose avec les big-data d'un outil extraordinaire, il faut revenir aux méthodes classiques de l'expertise afin, dans des systèmes complexes, de repérer les facteurs en jeu, au lieu de se contenter de corrélation comme si on avait l'explicitation causale.*

**Kenneth Cukier** : Nous pouvons faire avec les *big-data* des choses que nous ne pouvions pas faire auparavant. Il est vrai que dans beaucoup de cas complexes, il nous faudra aussi des études à caractère causal, mais cela suppose d'avoir une hypothèse de départ. Or les *big-data* nous donnent la possibilité de repérer la bonne hypothèse, ce qui n'était pas le cas quand nous n'avions que peu de données, ou face à un phénomène que nous n'arrivons pas du tout à comprendre. Pensez à toutes les nouvelles questions auxquelles nous pouvons répondre si nous avons cinquante, cent ou mille fois plus de points de données ! Nous pouvons ainsi utiliser les corrélations pour poser de nouvelles questions, mais aussi pour avancer dans les questions que nous nous posons, même si à la toute fin il nous faudra peut-être une étude causale, qui ne sera efficace que parce que, grâce aux données, nous avons construit la bonne hypothèse.

**Dominique Leglu** : *Il s'agit d'une autre façon d'utiliser les données. Puisque nous parlons des conducteurs d'automobile,*

---

*vous mentionnez dans votre livre le cas très intéressant de Shigeomi Koshimizu qui montre que nous avons chacun une façon différente de nous asseoir, en raison de notre poids ou de notre posture. Grâce à des capteurs, on peut ainsi caractériser la façon dont une personne s'assoit au volant, et en tirer différentes applications. Par exemple, un système d'identification pour lutter contre le vol, ou d'alerte en cas d'assoupissement. Nous pouvons donc acquérir de l'information en vue de renforcer la sécurité – mais il faut prendre les décisions, arbitrer entre liberté du conducteur et sécurité.*

**Pascale Leclercq :** *Vous avez employé au début de votre intervention une expression qui m'a frappée, celle d'un « esprit Big-Data ». On voit bien en effet que l'apparition des big data bouscule un certain nombre de rationalités. Comment qualifieriez-vous cet « esprit Big-Data » ?*

**Kenneth Cukier :** L'expression m'est venue pendant la rédaction du livre, quand j'ai compris qu'au cœur du *big-data*, il y a l'idée que tout ce que nous faisons, même de plus banal, contient de la valeur. Prenons quelque chose d'aussi ordinaire que le fait de s'asseoir. Il y a cinq ans, il aurait semblé ridicule d'y voir une source d'informations, susceptible d'être mise sous forme de données ayant de la valeur. Mais grâce à l'« esprit *Big-Data* », c'est désormais le cas, et la manière de s'asseoir permet de concevoir un système d'anti-vol ou d'alerte en cas d'assoupissement du conducteur. L'« esprit *Big-Data* », c'est voir partout des possibilités de changer la façon dont on travaille. C'est exactement comme au début d'Internet : des gens en marge des

---

grandes entreprises ont « l'esprit internet » et imaginent de nouvelles façons de faire, par exemple de vendre des livres, ce qui donne Amazon, maintenant une des plus puissantes entreprises mondiales. La France n'a pas vraiment eu cet « esprit internet », mais, avec son niveau d'éducation très élevé elle a tout à fait la possibilité d'être à la pointe des *big-data*, pourvu que les Français commencent à regarder le monde avec l'esprit qu'il faut. Il y a toutefois une différence avec internet : dans le cas d'internet, les start-ups avaient l'avantage, elles n'étaient pas entravées par l'ancienne façon de penser. Or aujourd'hui, ce sont les sociétés déjà en place qui ont l'avantage, parce que ce sont elles qui recueillent les données.

**Dominique Leglu** : *Encore faut-il qu'elles les aient repérées ! Au sujet des billets d'avion, ce ne sont pas les grandes compagnies qui ont créé Farecast, c'est un « maverick » qui est allé chercher les données...*

**Kenneth Cukier** : Certes, mais c'est parce qu'il a eu la chance d'y avoir accès ! Les compagnies aériennes savent maintenant qu'elles doivent se réserver l'accès à leurs données, sans savoir vraiment pour l'instant ce qu'il faut en faire. Car d'un côté un système comme *Farecast* diminue leur bénéfice, mais de l'autre ce peut être une façon d'élargir le marché...

**Dominique Leglu** : *Il s'agit en effet d'une question stratégique essentielle aujourd'hui pour les géants américains comme Amazon, Google, Apple, Microsoft ou Twitter. Ce sont*

---

*ces entreprises qui détiennent actuellement beaucoup de données, et qui ont le choix entre les utiliser elles-mêmes, les vendre à d'autres sociétés, s'allier avec elles ou leur suggérer les usages possibles qu'elles peuvent faire des big-data.*

**Patrick Mallaé :** *Il me semble, contrairement à la première question, que dans l'expression « big-data », c'est le « big » qui pose problème, plus que les « data ». Nous avons toujours produit des connaissances qui se sont avérées fausses et qui ont évolué avec le temps, et le Big-Data n'y échappera pas. Ce qui me semble beaucoup plus important dans votre exposé est qu'au moyen des big-data il semble possible de repenser et de remodeler les comportements de la population. Dans tous les exemples que vous avez donnés, on voit que le Big-Data permet de valoriser un comportement jugé plus adapté. Ma crainte n'est donc pas tant le Big Brother qu'une forme de « Big Father » : un paternalisme dans lequel nous serions amenés nous-mêmes, par le biais des données offertes par nos appareils mobiles à modifier nos comportements en vue d'atteindre tel ou tel objectif. Dans le domaine de la santé, il peut s'agir d'aller mieux, de nous soigner, et l'objectif est peut-être moins sujet à controverse, mais dans bon nombre d'activités, c'est la firme qui fixe cet objectif. Le Big-Data pose peut-être un problème au niveau de la connaissance, mais c'est surtout, à mon sens, son utilisation en vue de produire de « bons » comportements qui constitue le principal problème.*

**Kenneth Cukier :** Je suis d'accord : si j'ai bien compris, le risque est que nous modifions nous-mêmes nos comportements en raison d'une surveillance constante. C'est même plus qu'un risque, c'est une réalité, et le but du livre que j'ai écrit avec Viktor Mayer-Schönberger

---

est de lancer le débat, notamment sur ces questions-là. Il ne faut pas oublier cependant que l'interaction entre *big-data* et comportement est un peu plus complexe. Car si nous changeons nos comportements, les algorithmes doivent alors être modifiés à leur tour. C'est l'un des problèmes de *Google Flu Trends* : lorsque l'épidémie devient publique, passe aux informations, les individus font des requêtes simplement parce qu'ils veulent suivre l'actualité, ce qui modifie le sens des requêtes. En ce qui concerne les risques posés par les *big-data*, je terminerai en disant qu'il y a, pour reprendre l'expression de Donald Rumsfeld, de « l'inconnu inconnu » : des risques que nous n'imaginons même pas. Et cependant, tout en sachant qu'il y aura des drames liés au *big-data*, nous devons essayer de nous y préparer.



---

## POSTFACE

Dans le tableau qui précède, l'enthousiasme pour l'innovation, le goût pour l'invention, semblent toujours l'emporter sur les inhibitions dues à l'inquiétude. Aucune crainte d'apocalypse n'est exprimée ; on n'y décèle nulle tonalité de catastrophisme, à rebours de l'idéologie dominante en Occident.

On ne peut cependant envisager sans vertige le type de surveillance *a priori* qu'exercent déjà les États (démocratiques ou non...) ainsi que les grandes multinationales sur nos comportements ; ces dernières étant souvent encore bien plus intrusives que les premiers... Les géants du Web peuvent déjà traiter et analyser nos données (navigation, géolocalisation, ciblage comportemental) ; il n'est pas nécessaire de faire preuve d'une grande perspicacité pour en deviner l'usage à court ou moyen terme en matière de santé, d'assurance, de finance, de transport, de communication, de marketing...

Hier précisément ciblée sur des individus, la surveillance s'est aujourd'hui transformée en un contrôle massif et généralisé à l'échelle planétaire, guerre contre le terrorisme aidant. Le débat fait rage quant aux menaces qui pèsent sur nos libertés individuelles, à la suite de la révélation des scandales successifs de l'espionnage auquel se livre le gouvernement américain (mais il n'est pas le seul) de toutes nos données.

---

La personne que chacun d'entre nous est appelé à être et à devenir pour son propre compte subit l'injonction permanente d'aligner ses pensées et ses actes sur la logique qui règne dans la sphère des marchandises. L'objectif qui peut paraître paradoxal est que chacun prenne l'initiative de faire librement ce qu'on attend de lui. Ce que les penseurs des années soixante du siècle dernier appelaient pour la critiquer la « société de consommation » désigne aujourd'hui non plus une vue abstraite de l'avenir économiquement déterminé dans le cadre d'une philosophie de l'histoire, mais une réalité très actuelle et très concrète. Cette réalité impose d'autant plus aisément ses contraintes qu'elle se présente comme toujours souriante, parfois un tantinet grivoise et, somme toute, assez généreuse. Le consommateur crédule contribue à son propre asservissement par insouciance. Il livre ses données sans difficulté. Que cette société se révèle être également une « société du spectacle » confère aux textes austères et parfois obscurs de Guy Debord, une manière de valeur prophétique renouvelée. Ce qui ne lui aurait pas déplu.

Avec l'irruption, à première vue purement technique, du *Big-Data*, ces données de masse (personnelles ou non), considérées comme la matière première de demain, nous allons, cette fois-ci pour de vrai, goûter aux délices ambigus d'une manière d'être nouvelle traversant et redessinant tous nos modes de vie.

Cette transformation très profonde touche ce qu'on appelle l'« intimité » de chacun d'entre nous. Entendons-

---

nous bien, la « vie privée » a depuis belle lurette fait l'objet d'études juridiques précises et de décisions politiques importantes <sup>6</sup>. Mais les interrogations sur le *Big-Data* débordent ce champ de réflexion fondamentalement, sinon étroitement, juridique.

La vie intime signale l'existence d'une réalité sous-jacente à la vie intérieure. L'intime désigne le plus personnel d'un être, le plus secret. Ce secret hautement valorisé, doit être caché et inaccessible aux autres. L'intime, plus que l'intérieur, suppose une intensité qui se traduit par des liens spécifiques avec autrui. Ne dit-on pas que nos « amis intimes » sont des amis plus solides que tous autres ? Quant à nos relations intimes, chacun sait ce qu'il en est pour soi-même.

L'espoir ne nous quitte jamais d'exercer quelque contrôle sur ce qui nous arrive. Nos amours, nos déceptions (sentimentales ou autres), nos ambitions et même nos colères se trouvent soumis à un examen sans fin, souvent non sans quelque complaisance. On sait que les êtres humains « extériorisent » plus ou moins. Les psychologues en ont même fait une catégorie savante qu'ils ont baptisé l'extraversion. Laquelle laisse sans voix l'introverti qui nourrit à l'encontre de l'extraverti un soupçon de superficialité. Les valeurs des recruteurs et des DRH ne sont pas universellement partagées...

---

6. Voir notamment les carnets publiés par l'Institut Diderot téléchargeables sur [www.institutdiderot.fr](http://www.institutdiderot.fr) : *L'avenir de la cybersécurité* de Nicolas Arpagian, *L'avenir de la vie privée dans la société numérique* d'Alex Türk et *L'avenir de la propriété intellectuelle* de Denis Olivennes.



---

Que nous annoncent les oracles de nos ingénieurs quelque peu éberlués, sinon, pour certains, ivres de leur propre puissance analytique ? Pour demain, un incomparable pouvoir ! On l'entend à mi-mots tout au long de leurs discours prospectifs sur le *Big-Data*. Demain, ils devanceront nos propres désirs de citoyen-consommateur. Le rationalisme appliqué nous ramènera étrangement sur la voie de la pensée magique ! Une puissance anonyme et inconnue, forte de l'étendue des connaissances acquises sur vos comportements, prétendra désormais abolir, ou du moins contourner, la distance et le temps qui vous sépare de la réalisation de vos désirs. Ce sera, c'est déjà, la société d'impatience !

Pourquoi ce qui vaut pour l'achat d'un livre, ne vaudrait-il pas, par exemple pour celui d'un appartement ou d'un programme des prochains candidat(e)s à la présidence ? Pourquoi nos objets d'amour y échapperaient-ils ? Internet place notre vie sous le signe de la rencontre, dans tous les domaines de l'existence. Les sociologues nous alertent sur les conséquences humaines du rapprochement de la rencontre amoureuse et de la rencontre commerciale. Il existe des sites pour tout. De l'ordre est mis dans nos relations par les milliards d'internautes amoureux sous l'impulsion et la direction d'un nombre restreint de grandes entreprises multinationales. Le concept même d'être « amoureux » doit être revu à l'expérience d'Internet. Les interrogations se bousculent. Les couples survivront-ils à la sérialité, en droit infinie, des coups de foudre ? Le marché du mariage, tel qu'il a été institué,

---

va-t-il « tenir » devant l'expansion de celui du divorce ? Le désir de tout savoir de l'autre gardera-t-il quelque attrait si autrui dans sa singularité m'est offert dans sa nudité la plus immédiate ? Ma rencontre peut-elle être conçue comme un « plan » dicté par une instance transcendante paradoxalement omniprésente dans notre monde ? Nous ne sommes aujourd'hui sûrs de rien. Nous en sommes au point de devoir parier en faveur de ce nouveau monde, sauf à nous exclure de la connexion universelle et très bientôt de l'ensemble de la société.

Reste que la transformation digitale qui me permettra, après la gestion de ma vie intime, de me faire soigner à la suite d'une déception ou de souscrire une assurance pour rouler dans une Apple-car peut laisser entrevoir une sorte de synergie terrifiante à la manière du Skynet, l'intelligence artificielle antagoniste de la série *Terminator*. La question n'est pas aujourd'hui de libérer l'homme de la machine, mais de libérer les machines des usages auxquels nous les avons asservies. Il se pourrait alors que l'humanité y retrouve une part du bien dont elle se trouve effectivement spoliée.

Dominique Lecourt

Directeur général de l'Institut Diderot

---

Retrouvez l'intégralité du débat en vidéo sur  
[www.institutdiderot.fr](http://www.institutdiderot.fr)

---

## LES PUBLICATIONS DE L'INSTITUT DIDEROT

### Dans la même collection

- L'avenir de l'automobile - Louis Schweitzer
- Les nanotechnologies & l'avenir de l'homme - Etienne Klein
- L'avenir de la croissance - Bernard Stiegler
- L'avenir de la régénération cérébrale - Alain Prochiantz
- L'avenir de l'Europe - Franck Debié
- L'avenir de la cybersécurité - Nicolas Arpagian
- L'avenir de la population française - François Héran
- L'avenir de la cancérologie - François Goldwasser
- L'avenir de la prédiction - Henri Atlan
- L'avenir de l'aménagement des territoires - Jérôme Monod
- L'avenir de la démocratie - Dominique Schnapper
- L'avenir du capitalisme - Bernard Maris
- L'avenir de la dépendance - Florence Lustman
- L'avenir de l'alimentation - Marion Guillou
- L'avenir des humanités - Jean-François Pradeau
- L'avenir des villes - Thierry Paquot
- L'avenir du droit international - Monique Chemillier-Gendreau
- L'avenir de la famille - Boris Cyrulnik
- L'avenir du populisme - Dominique Reynié
- L'avenir de la puissance chinoise - Jean-Luc Domenach
- L'avenir de l'économie sociale - Jean-Claude Seys
- L'avenir de la vie privée dans la société numérique - Alex Türk
- L'avenir de l'hôpital public - Bernard Granger
- L'avenir de la guerre - Henri Bentegeat & Rony Brauman
- L'avenir de la politique industrielle française - Louis Gallois
- L'avenir de la politique énergétique française - Pierre Papon
- L'avenir du pétrole - Claude Mandil
- L'avenir de l'euro et de la BCE - Henri Guaino & Denis Kessler
- L'avenir de la propriété intellectuelle - Denis Olivennes
- L'avenir du travail - Dominique Méda

- 
- L'avenir de l'anti-science - Alexandre Moatti
  - L'avenir du logement - Olivier Mitterand
  - L'avenir de la mondialisation - Jean-Pierre Chevènement
  - L'avenir de la lutte contre la pauvreté - François Chérèque
  - L'avenir du climat - Jean Jouzel
  - L'avenir de la nouvelle Russie - Alexandre Adler
  - L'avenir de la politique - Alain Juppé

## Les Notes de l'Institut Diderot

- L'euthanasie, à travers le cas de Vincent Humbert - Emmanuel Halais
- Le futur de la procréation - Pascal Nouvel
- La République à l'épreuve du communautarisme - Eric Keslassy
- Proposition pour la Chine - Pierre-Louis Ménard
- L'habitat en utopie - Thierry Paquot
- Une Assemblée nationale plus représentative - Eric Keslassy
- Où va l'Égypte ? - Ismaïl Serageldin
- Sur le service civique - Jean-Pierre Gualazzi
- La recherche en France et en Allemagne - Michèle Vallenthini
- Le fanatisme - Texte d'Alexandre Deleyre  
présenté par Dominique Lecourt

## Les Dîners de l'Institut Diderot

- La Prospective, de demain à aujourd'hui - Nathalie Kosciusko-Morizet
- Politique de santé : répondre aux défis de demain - Claude Evin
- La réforme de la santé aux Etats-Unis :  
quels enseignements pour l'assurance maladie française ? - Victor Rodwin
- La question du médicament - Philippe Even

## Les Entretiens de l'Institut Diderot

- L'avenir du progrès (actes des Entretiens 2011)

# L'avenir des Big-Data

L'enthousiasme pour l'innovation et le goût pour l'invention semblent toujours l'emporter sur les inhibitions dues à l'inquiétude, devant le tableau dressé du Big-Data, les données de masse (personnelles ou non) considérées comme la matière première de demain.

Avec l'irruption, à première vue purement technique du Big-Data, nous allons goûter aux délices ambigus d'une manière d'être nouvelle traversant et redessinant tous nos modes de vie.

On ne peut cependant envisager sans vertige le type de surveillance a priori qu'exercent déjà les États (démocratiques ou non...) ainsi que les grandes multinationales sur nos comportements ; ces dernières étant souvent encore bien plus intrusives que les premiers... Les géants du Web peuvent déjà traiter et analyser nos données (navigation, géolocalisation, ciblage comportemental), il n'est pas nécessaire de faire preuve d'une grande perspicacité pour en deviner l'usage à court ou moyen terme en matière de santé, d'assurance, de finance, de transport, de communication, de marketing...

Hier précisément ciblée sur des individus, la surveillance s'est aujourd'hui transformée en un contrôle massif et généralisé à l'échelle planétaire, guerre contre le terrorisme aidant. Le débat fait rage quant aux menaces qui pèsent sur nos libertés individuelles.

**Dominique LECOURT**

Directeur général de l'Institut Diderot



“ Kenneth CUKIER est éditeur « data » au magazine *The Economist* ; il collabore aussi au *New-York Times*, *Financial Times* et au *Foreign Affairs*.



“ Dominique LEGLU est directrice de *Sciences et Avenir*

La présente publication ne peut être vendue