



VIVANT MATIÈRE SOCIÉTÉS UNIVERS TERRE

NUMÉRIQUE MES THÈMES

Types



DOSSIER

Paru le 15.11.2012
Mis à jour le 29.01.2014



Big Data, la déferlante des octets

BILLET

 TERRE (/TERRE)  NUMÉRIQUE (/NUMERIQUE)

LES INÉDITS DU CNRS (/LES-INEDITS-DU-CNRS)

Le Big Data est-il polluant ?

02.04.2015, par

Grandes conso



Partager l'article

technologies de l'information ont un coût environnemental qui n'a rien de virtuel.

Spécialistes en informatique, Jean-Marc Pierson et Laurent Lefèvre nous expliquent en quoi consiste cette pollution.

Une fois par mois, retrouvez sur notre site les Inédits du CNRS, des analyses scientifiques originales publiées en partenariat avec Libération. (<http://www.liberation.fr/>)



Du commerce à l'énergie en passant par la finance, la santé, le transport, la culture et même la science, les données numériques représentent pour beaucoup le pétrole du XXI^e siècle. Chaque jour, l'industrie du Big Data se fait plus grosse, représentant déjà plus de 4 millions d'emplois directs dans le monde ! Mais il y a une donnée que nous avons peut-être oubliée en route : tout a un coût. Et celui du Big Data est écologique. Derrière cette informatique virtualisée, distribuée et distante, se trouvent, en effet, des infrastructures bien réelles qui ont une consommation énergétique et un impact carbone forts. N'ayons pas peur des mots : oui, le Big Data pollue.

Des infrastructures pharaoniques

Mais reprenons du début. Le Big Data (ou « mégadonnées » en français) désigne tout à la fois la faculté de produire ou de collecter des données numériques, de les stocker, de les analyser et de les visualiser. Il est très souvent défini par ses caractéristiques liées aux « 3V » (volume, variabilité, vitesse) : les données arrivent en masse, notamment avec l'arrivée conjointe de l'Internet des objets, à une vitesse sans précédent et sont de nature plus variées que par le passé. En 2015, le patrimoine mondial de données atteindra 8 zétaoctets (10^{21} octets). Des infrastructures pharaoniques sont donc déjà nécessaires pour stocker cette avalanche de données, mais il faut y ajouter celles pour les traiter. L'analyse de données, provenant

d'observations de l'environnement d'expériences scientifiques ou encore de données m
 importants, concentrés dans des grands centres et des supercalculateurs.



Partager l'article



Dans son data center, Google doit conserver la trace de milliards de pages Web. 🔄

(#image-aid-293)

© GOOGLE

Prenons l'exemple de Google, qui indexe des millions de documents pour faciliter et accélérer la recherche. Ces données sont aussi analysées pour pouvoir proposer aux utilisateurs des contenus publicitaires : c'est sur cette approche que repose son modèle économique. Or, pour réaliser ces traitements, on estime que Google possède plus d'un million de serveurs, comme les trois autres géants du Web que sont Amazon, Microsoft et Facebook. Qu'elles possèdent leurs propres infrastructures ou qu'elles les louent, les entreprises et les institutions plus modestes font souvent appel à des milliers d'équipements interconnectés, depuis la capture de données jusqu'à leur analyse.

Une demande énergétique croissante à contenir

Des travaux, comme ceux du groupement de service Écoinfo (<http://ecoinfo.cnrs.fr>), créé par le CNRS, sont venus rappeler à quel point les technologies de l'information sont consommatrices d'énergie et produisent des gaz à effet de serre à toutes les étapes de leur cycle de vie : conception et transport des équipements, usage et fin de vie. En phase d'utilisation, les éléments essentiels du Big Data peuvent se répartir en trois catégories : les équipements terminaux, les réseaux et

les centres de données qui consomment chacune une puissance électrique comparable en 2013, soit une quarantaine de tranches de centrales nucléaires^[1] (#footnote1_9pm06gy). Un chiffre qui a forcément des répercussions sur le climat, même si cet impact carbone dépend du mix énergétique du pays d'usage (34 grammes de CO₂ par kwh en France en février).



Partager l'article

“Les technologies de l’information produisent des gaz à effet de serre à toutes les étapes de leur cycle de vie.”

Alors, en arriverons-nous un jour au scénario (provocateur) de la société Cisco, dans lequel seules certaines machines auront le droit de communiquer (celles qui ont une adresse IP paire par exemple), à l’instar de la circulation alternée utilisant les plaques minéralogiques des voitures ? Fort heureusement, une multitude de réponses et d’alternatives se mettent en place. Des codes de conduites sont proposés aux industriels et aux hébergeurs pour améliorer leurs infrastructures de calcul, de communication et de stockage à grande échelle. Des innovations sont appliquées sur le design des équipements et leur potentialité de recyclage. Des recherches sur l’efficacité énergétique proposent de nouvelles voies pour limiter ces impacts. Autant d’initiatives indispensables si l’on veut pouvoir contenir la demande énergétique croissante de ces infrastructures tout en garantissant une bonne qualité de service aux utilisateurs.

Une aide pour réduire d’autres sources de pollution ?

Si le Big Data, en lui-même, est polluant, permet-il, par ailleurs, de diminuer d’autres sources de pollution ? Plusieurs exemples semblent en témoigner. Ainsi, l’analyse de données massives peut permettre d’optimiser des processus industriels et donc de réduire les émissions polluantes associées. De même, les fermiers peuvent recevoir en temps réel des informations sur leurs plantations, issues de capteurs et d’images satellites, afin de ne diffuser que la quantité d’eau nécessaire et le juste dosage de pesticides.

“L’analyse de données permet de réduire les émissions polluantes industrielles.”

[Partager l'article](#)

La simple collecte des données et leur mise à disposition peuvent également amener des changements de comportements. À Portland, aux États-Unis, une communauté de citoyens déploie des capteurs mesurant la qualité de l'air dans leur voisinage. Les données, de meilleure précision et qualité que les données mesurées par l'Agence de protection de l'environnement, sont analysées et visualisées sur un site Web. La présence de ces données ouvertes amènera de nouveaux usages et comportements, moins centrés sur les voitures en cas de pic de pollution. Enfin, un dernier exemple vient de Chine où la ville de Pékin s'est lancée, avec IBM, dans un programme de diminution de la pollution en ville (dont on connaît l'ampleur). Pour cela, des données sont récoltées par une multitude de capteurs, couplés avec des données satellites de météorologie, et analysées par des systèmes d'intelligence artificielle à grande échelle. Le but : produire une carte de prédiction de la pollution 72 heures en avance.

Le Big Data est la ruée vers l'or des temps modernes. Comme sa glorieuse aînée, elle draine beaucoup d'espairs – fondés ou non – et pose d'importants problèmes tout en permettant le développement de nouveaux territoires. Il nous appartient de veiller à ce que le coût environnemental de ces technologies soit contrebalancé – au moins en partie – par des progrès dans la lutte contre le réchauffement climatique et la pollution.

Notes

1. (#foundry) Une tranche est une unité de production électrique au sein d'une centrale.
-



[\(/billets/oui-les-musulmans-sont-en-accord-avec-la-republique\)](/billets/oui-les-musulmans-sont-en-accord-avec-la-republique)

 **SOCIÉTÉS** (/SOCIETES)

« Oui, les musulmans sont en accord avec la République » [\(/billets/oui-les-musulmans-sont-en-accord-avec-la-republique\)](/billets/oui-les-musulmans-sont-en-accord-avec-la-republique)
31.03.2015



[\(/billets/la-cinquante-et-unieme-nuance-de-lautonomie\)](/billets/la-cinquante-et-unieme-nuance-de-lautonomie)

La cinquante-et-unième nuance de l'autonomie

◆ SOCIÉTÉS (/SOCIETES)



Partager l'article

(/billets/la-cinquante-et-unieme-09.03.2015-nuance-de-lautonomie)



(/billets/les-sciences-a-lage-du-libre-acces)

■ NUMÉRIQUE (/NUMERIQUE)

Les sciences à l'âge du libre accès (/billets /les-sciences-a-lage-du-libre-acces)

04.03.2015



Partager l'article

[\(/billets/cogitez-si-vous-voulez-les-decisions-sont-irrationnelles\)](#)

 **VIVANT** (/VIVANT)  **SOCIÉTÉS** (/SOCIETES)

Cogitez si vous voulez, les décisions sont
irrationnelles (/billets/cogitez-si-vous-voulez-

les-decisions-sont-irrationnelles)
05.03.2015