

Implication des Big Data en sciences de la vie

Publié le vendredi 30 janvier 2015

Voir en ligne : <https://www.france-science.org/Implication-des-Big-Data-en,4435.html>

La recherche dans le domaine de la santé ne cesse de progresser grâce aux avancées technologiques qui permettent des analyses plus rapides et de moins en moins coûteuses [1]. Cependant, cela génère des quantités de données de plus en plus importantes : séquençages, essais cliniques et *crowdsourcing* ne sont que quelques exemples d'applications à l'origine de l'augmentation exceptionnelle du volume de données biologiques. Ces ensembles de données deviennent tellement volumineux qu'ils deviennent difficiles à traiter avec des outils classiques de gestion de base de données ou de gestion de l'information ; ils sont communément appelés "Big Data" [2]. La gestion de ces mégadonnées entraîne des problèmes de protection des données, de confidentialité, voire de sécurité nationale.



Crédits : Kentoh

En décembre 2014, plusieurs experts du *Center for Science, Technology and Security Policy* (CSTSP) de l'*American Association for the Advancement of Science* (AAAS), la *Biological Countermeasures Unit* (BCU) du *Federal Bureau of Investigation* (FBI) *Weapons of Mass Destruction Directorate* (WMDD) et le *United Nations Interregional Crime and Justice Research Institute* (UNICRI) ont publié un rapport sur les implications des Big Data en sciences de la vie sur la sécurité nationale et internationale (*National and Transnational Security Implications of Big Data in the Life Sciences*) [3]. Ce rapport fait suite à une rencontre qui a eu lieu en avril 2014, dont les objectifs étaient les suivants :

- examiner les risques et les bénéfices associés aux analyses de données de masse ;
- développer des cadres d'évaluation et de gestion des risques et bénéfices des technologies émergentes telles que les Big Data en sciences de la vie ;
- identifier les moyens à mettre en oeuvre par le gouvernement américain pour caractériser davantage les risques et bénéfices des analyses Big Data et pour limiter les risques potentiels.

A l'ère des nouvelles technologies de l'information et du *cloud-computing*, l'implication des Big Data en sciences de la vie suscite l'inquiétude des autorités gouvernementales en raison de l'éventuelle vulnérabilité des ensembles de données et le détournement de ces informations pour le développement redouté d'armes de destruction massive de nature biologique. De nombreuses réglementations ont été mises en place aux Etats-Unis et à travers le monde pour empêcher que des agents biologiques ou des toxines ne soient développés à des fins terroristes. Malheureusement, ces réglementations ne s'appliquent pas à l'utilisation des données de masse et d'outils de calcul qui pourraient permettre de concevoir des armes biologiques (par exemple : exploitation de données de séquençage pour concevoir des petites molécules qui pourront interagir avec les protéines d'organismes vivants pour modifier leur comportement dans un sens déterminé). Un intérêt particulier doit donc être porté à la protection des données et restreindre leur utilisation. Des moyens de prévention et de détection de détournement doivent être développés. Les autorités gouvernementales et les communautés scientifiques sont des acteurs majeurs pour contribuer à leur mise en application et veiller à leur pérennité.

Les avancées technologiques et la recherche biomédicale progressent grâce aux Big Data, notamment dans les domaines suivants : médecine personnalisée et développement de médicaments, prestations de soins en santé ou encore préservation des écosystèmes face au changement climatique. Cependant, des efforts restent à faire dans le développement des puissances de calcul, le stockage des données, l'accès et les équipements. Avant de pouvoir procéder à des analyses en masse, il est primordial d'établir des lignes

directrices telles que standardiser les terminologies, s'assurer de protéger la vie privée et la confidentialité des contributeurs et sécuriser l'accès aux données.

Les investissements d'agences gouvernementales, de fondations privées et de grands groupes sont de plus en plus importants et montrent bien l'intérêt grandissant pour ce domaine. En mars 2012, la Maison Blanche a annoncé un investissement de 200 millions de dollars pour la recherche et le développement en Big Data [4]. Suite à l'annonce de cette "*Big Data Initiative*", de nombreuses agences gouvernementales ont commencé à financer et apporter leur soutien aux recherches concernées. A titre d'exemple, les *National Institutes of Health* (NIH) ont créé la *Big Data to Knowledge* (BD2K) *Initiative* en 2012 dans le but de permettre à la fois le développement d'approches et de méthodes innovantes et de mettre en place des programmes éducatifs pour mieux exploiter les Big Data au profit de la recherche biomédicale [5, 6]. En octobre 2014, douze centres d'excellence à travers les Etats-Unis ont été nommés. Plusieurs subventions leur ont été allouées pour un total de 32 millions de dollars pour l'année fiscale 2015. Francis Collins, directeur des NIH, souhaite que ce consortium investisse 656 millions de dollars au cours de sept prochaines années pour collecter, analyser et organiser les résultats de recherche, les analyses génomiques et autres données de masse à mettre à disposition de la communauté scientifique. En Californie, quatre universités ont été sélectionnées pour accueillir des centres d'excellence. L'un d'eux se trouve à l'Université de Californie à Los Angeles (UCLA), en partenariat avec cinq autres instituts dont le *Scripps Research Institute* de La Jolla, destiné à générer des données sur les marqueurs protéiques des maladies cardiovasculaires, et deux centres, qui étudieront les maladies cérébrales, sont situés à l'Université de Californie du Sud (*University of Southern California* - USC) à Los Angeles.

Outre les études cliniques, les données de masse en santé sont également enrichies par les objets connectés développés par des géants de l'informatique tels que Google, Samsung ou encore Apple. Ces petits équipements gagnent en popularité auprès du grand public et permettent de recueillir des informations précieuses au quotidien.

Sources :

- [1] "Médecine individualisée : le génome à 1000 dollars, l'interprétation à 1 million..." - BE Etats-Unis 362 - AGENES Fabien - 21/03/2014 - <http://www.bulletins-electroniques.com/actualites/75480.htm>
- [2] "Big data" - Wikipedia - 08/01/2015 (dernière mise à jour) - http://fr.wikipedia.org/wiki/Big_data
- [3] "National and Transnational Security Implications of Big Data in the Life Sciences" - AAAS [en ligne] AAAS/CSTSP, FBI/WMD/BCU, UNICRI - 11/12/2014 - <http://www.aaas.org/report/national-and-transnational-security-implications-big-data-life-sciences>
- [4] "Obama Administration Unveils "Big Data" Initiative : Announces \$200 Million in New R&D Investments" - White House Office of Science and Technology Policy - Executive Office of the President - 29/03/2012 - http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf
- [5] "NIH Big Data to Knowledge (BD2K)" - <http://bd2k.nih.gov/index.html#sthash.XHkn4Wtf.dpbs>
- [6] "Big data, meet big money : NIH funds centers to crunch health data" - LA Times [en ligne] HEALY Melissa - 09/10/2014 - <http://www.latimes.com/science/sciencenow/la-sci-sn-big-data-money-20141009-story.html>

Pour en savoir plus, contacts :

- "Conférence "Informed Health 2015" à UCSF : collecter et traiter les données sur les patients pour changer notre rapport à la santé" - BE Etats-Unis 390 - LOURDANI Hocine - 23/01/2015 - <http://www.bulletins-electroniques.com/actualites/77682.htm>
- "De nombreux partenariats d'envergure autour du séquençage génétique" - BE Etats-Unis 390 - LOURDANI Hocine - 23/01/2015 - <http://www.bulletins-electroniques.com/actualites/77680.htm>
- "Californie : un partenariat prometteur dans la lutte contre le cancer" - BE Etats-Unis 388 - METIVIER Clément - 09/01/2015 - <http://www.bulletins-electroniques.com/actualites/77620.htm>
- "Le Conseil Atlantique sur la cybersécurité : la collaboration contre les menaces peut être la clé de la prospérité" - BE Etats-Unis 374 - ZAMSKY Ellen - 20/06/2014 - <http://www.bulletins-electroniques.com/actualites/76202.htm>

- "La Maison Blanche s'intéresse à la menace du Big Data sur la vie privée de ses citoyens (partie 1/2)" - BE Etats-Unis 370 - HUYNH Maxime, FENIEYS Guilhem - 23/05/2014 - <http://www.bulletins-electroniques.com/actualites/75992.htm>

- "La Maison Blanche s'intéresse à la menace du Big Data sur la vie privée de ses citoyens (partie 2/2)" - BE Etats-Unis 371 - DEBACKER Thomas - 30/05/2014 - <http://www.bulletins-electroniques.com/actualites/76035.htm>

- "Les nouvelles technologies au service de la réforme du système de santé américain" - BE Etats-Unis 359 - BOUQUET Basile, POISSON Thibault - 28/02/2014 - <http://www.bulletins-electroniques.com/actualites/75301.htm>

Code brève

ADIT : 77730

Rédacteurs :

- Viviane Chansavang, Attachée scientifique adjointe - Los Angeles, deputy-sdv.la@ambascience-usa.org ;
- Retrouvez toutes nos activités sur le site du Consulat général de France à Los Angeles - <http://www.consulfrance-losangeles.org/spip.php?rubrique241> ;
- Retrouvez toutes nos activités sur le site de la MST de l'Ambassade de France aux Etats-Unis : <http://france-science.org>.