

# Une science plus ouverte

Le rédacteur en chef de la revue *Nature* et un dirigeant de la Royal Society discutent, lors d'une rencontre à la FAPESP, des enjeux et les limites de l'ouverture des données scientifiques

**Bruno de Pierro**

PUBLIÉ EN MARS 2013

« L'ouverture des données en soi n'a aucune valeur, car une science ouverte c'est beaucoup plus que la simple mise à disposition de données scientifiques », estime le physicien anglais Philip Campbell, rédacteur en chef de la revue *Nature*, l'une des revues scientifiques les plus prestigieuses au monde. De passage à São Paulo, Philip Campbell a participé à la rencontre Science as an Open Enterprise: Open Data for Open Science, réalisée le 25 février à la FAPESP, où, dans un auditorium rempli, il a parlé des enjeux et des transformations liés à l'ouverture des données scientifiques. Durant l'évènement, le directeur scientifique de la Fondation, Carlos Henrique de Brito Cruz, a parlé des perspectives d'ouverture des données scientifiques au Brésil. Le président-directeur du Conseil Technico-Administratif de la FAPESP, José Arana Varela, le Secrétaire des Relations Extérieures de la Royal Society, Martyn Poliakoff, ont participé en tant que médiateurs.

Les discussions ont tourné autour d'un rapport publié en juin 2012 par la Royal Society. Dans ce rapport, la plus ancienne société scientifique du monde souligne la nécessité d'appréhender l'ouverture de données scientifiques chaque fois plus nombreuses, mais dont le contenu n'est pas toujours intelligible ou de l'intérêt des chercheurs. « De rapides changements technologiques ont créé de nouvelles formes d'acquisition, de stockage, de manipulation et de transmission d'ensembles de données qui favorisent de nouveaux moyens de communication et de collaboration », déclare Martyn Poliakoff. L'étude fait suite à une polémique

déclenchée en 2009, au Royaume-Uni, sur des messages e-mails envoyés par des climatologues qui ont été piratés et publiés. Les messages suggéraient qu'un scientifique avait essayé de cacher des données contraires aux évidences du réchauffement planétaire. Une enquête a écarté l'hypothèse de fraude mais l'affaire a déclenché des débats sur la nécessité d'avoir une science plus ouverte.

Philip Campbell explique que les données scientifiques ne doivent pas être simplement accessibles mais qu'elles doivent être traitées pour devenir compréhensibles et réutilisables. Le physicien souligne que l'une des raisons qui font de l'ouverture des données un thème qui ne peut être remis à plus tard est que la science pourrait gagner davantage notre confiance grâce à la réplication et à la reproductibilité des données de recherche. Selon lui, cela permettrait de lutter contre les fraudes dans le monde universitaire et d'élargir la participation publique dans la science.

Un exemple de collaboration publique est présenté dans le rapport de la Royal Society. En 2011, un foyer d'infection intestinale provoquée par *Escherichia coli* est apparu en Allemagne et s'est répandu en Europe, affectant environ 400 mille personnes. Les médecins hambourgeois n'arrivaient pas à trouver une solution, car, à première vue, la bactérie était identique à celles d'autres souches. Le problème n'a été résolu qu'après que les données sur le génome de la souche de *E. coli* soient devenues accessibles et publiées sur un site à la portée de n'importe quel chercheur. Peu de temps après, environ 200 rapports scientifiques



## La technologie a créé de nouveaux moyens de communication et de collaboration scientifique, comme le montre le rapport

Philip Campbell, de la revue *Nature* (à gauche), et Martyn Poliakoff, de la Royal Society; l'ouverture peut augmenter la participation publique dans la science

PHOTOGRAPHIES LEO RAMOS

ont été publiés et ont indiqués ce qui pourrait être fait pour stopper l'épidémie.

«De nos jours, l'information détermine comment les scientifiques doivent s'adapter aux changements technologiques, sociaux et politiques, avec de profondes implications dans la manière comment la science est conduite et communiquée», déclare Martyn Poliakoff. Le chimiste a également parlé d'un projet auquel il participe et appelé *PeriodicVideos* ([www.periodicvideos.com](http://www.periodicvideos.com)), qui propose des vidéos ludiques de diffusion scientifique. Le projet a commencé en 2008, à partir d'un partenariat entre l'Université de Nottingham et la BBC de Londres.

### TROIS MODÈLES

Un autre thème discuté au cours de cette rencontre, outre la publication des données produites par la recherche, a concerné l'élargissement des outils utilisés pour avoir accès aux articles publiés dans les revues scientifiques. Campbell a rappelé qu'il existe aujourd'hui trois modèles pour la publication ouverte d'articles. Le premier met librement l'article à disposition au bout d'une période maximum de 6 à 12 mois après sa publication. Dans le second, l'article peut être consulté au moment de sa publication et le troisième modèle est une version hybride, qui ne met librement à disposition qu'une partie du contenu si

l'auteur paie une taxe pour une diffusion immédiate. Questionné sur les effets que pourraient causer l'ouverture de données sur la communication scientifique, principalement pour le journalisme d'information scientifique, Philip Campbell ne croit pas que l'ouverture améliorera forcément le processus de communication. D'après lui, peu importe si les articles sont librement accessibles ou mis à disposition au moyen du paiement d'une taxe, car les bonnes revues évalueront toujours leurs résultats et ne publieront que les meilleurs. Ce qui peut se passer, selon lui, c'est que le feedback peut être plus rapide dans une revue qui fournit un contenu ouvert grâce aux corrections ou aux commentaires apportés dès la publication. Ceci peut dans une certaine mesure améliorer l'article. «Le contenu librement ouvert peut être préférable dans ce cas là (par rapport à celui qui est accessible après le paiement d'une taxe)», déclare-t-il.

Le Directeur Scientifique de la FAPESP, Carlos Henrique de Brito Cruz, a rappelé que la discussion autour des données ouvertes n'était pas complètement nouvelle dans le pays, en citant des banques de données brésiliennes consultables sur Internet. Il a mentionné les informations ouvertes fournies par l'Institut National de Recherches Spatiales (Inpe), l'Institut Brésilien de Géographie et de Statistique

(IBGE) et la Plateforme Lattes, du Conseil National de Développement Scientifique et Technologique (CNPq), qui fournit des informations universitaires provenant de plus de 174 mille chercheurs. Brito Cruz a également souligné l'importance de la bibliothèque virtuelle SciELO, de la FAPESP, qui permet d'avoir librement accès à 270 revues nationales avec environ 1 million de consultations quotidiennes. «C'est un mécanisme important qui a permis d'apporter davantage de visibilité à la science brésilienne dans le monde», déclare-t-il. L'autre avancée importante a été un accord signé il y a deux ans pour la création d'un répertoire central de tous les articles liés à la recherche qui ont reçu un financement de la FAPESP, et qui deviendront accessibles en respectant les normes de la revue scientifique qui aura publié chaque article. Ce répertoire devrait être prêt d'ici la fin de l'année.

En ce qui concerne le Brésil, Philip Campbell dit ne pas connaître le nouvel agenda scientifique, mais il reconnaît les efforts entrepris par le Brésil pour se positionner sur le plan international. Il cite comme exemple le soutien apporté par la FAPESP aux projets de recherche qui associent les scientifiques et les entreprises. Je pense que ceci est un type de financement précieux, outre l'élargissement des relations entre les universités et l'industrie». ■