



AEF Dépêche n°559161 - Paris, le 22/03/2017 11:48:00
- Recherche et Innovation -

- 217.114.201.5 - www.aef.info

Toute reproduction ou transmission de cette dépêche est strictement interdite, sauf accord formel d'AEF.

TDM : "Le métier du chercheur change, il faut savoir l'évaluer de façon différente" (Olivier Le Gall, Inra, RUE 2017)

Par **René-Luc Bénichou**



Olivier Le Gall, directeur général délégué aux affaires scientifiques de l'Inra.

© Inra

La pratique du TDM amènera à "revoir le processus d'évaluation des chercheurs", prédit Olivier Le Gall, directeur général délégué aux affaires scientifiques de l'Inra. Il intervient lors d'une table ronde aux RUE, le 16 mars 2017, sur le thème : "Recherche et open data : la loi numérique et le TDM donneront-ils à la France un avantage majeur en termes d'open science ?" Pour lui, il paraît inéluctable qu'avec le TDM, "le métier du chercheur change" et qu'il faudra par conséquent "savoir le reconnaître et l'évaluer de façon différente". Claude Kirchner, conseiller du président d'Inria, et Renaud Fabre, directeur de l'information scientifique et technique du CNRS, confirment que le TDM ouvre "une nouvelle frontière pour le développement des activités" scientifiques, ainsi que l'illustrent à l'aide d'exemples concrets Claire Nédellec et Mathieu Andro, de l'Inra.



Table ronde sur le TDM aux RUE, le 16 mars 2017, avec de gauche à droite : Claude Kirchner (Inria), Claire Nédellec (Inra), Renaud Fabre (CNRS), Mathieu Andro (Inra) et Olivier Le Gall (Inra).

© Sébastien Ferraro / Groupe AEF

Le TDM (text and data mining), autrement dit la fouille de textes et de données, peut s'apparenter à un "téléscope numérique" permettant d'explorer la littérature scientifique – "le texte, mais aussi toutes les données d'expérimentation et d'usage" –, ce qui "donne la capacité de découvrir de nouvelles choses", décrit Claude Kirchner, conseiller du président d'Inria, lors d'une table ronde dédiée à ce sujet lors des RUE, le 16 mars 2017 à Paris.

RENAUD FABRE (CNRS) : CRÉER UNE AGENCE POUR ALLER PLUS LOIN

Après l'adoption de la loi pour une République numérique, qui accorde le droit aux chercheurs de pratiquer le TDM (lire sur AEF), le décret d'application est maintenant en cours d'examen au Conseil d'État, confirme Renaud Fabre, directeur de l'information scientifique et technique du CNRS. "Si nous n'avions pas cette loi numérique, nous aurions un inconvénient majeur car tous les grands pays, aujourd'hui, ont déjà initié des démarches qui mènent au TDM." Cela dit, précise-t-il, "seuls deux pays en Europe peuvent commencer à avancer" concrètement sur le sujet : l'Angleterre et la France", l'Europe dans son ensemble se situant plutôt "en retard".

Renaud Fabre prévient par ailleurs qu' "un seul texte dans un seul coin ne traitera pas toutes les questions qui restent ouvertes après la loi numérique". C'est la raison pour laquelle le CNRS propose "une analyse systémique des textes de la loi, qui doit donner lieu à un guide qu'il faudrait construire avec tous les acteurs" (lire sur AEF). "Nous appelons de nos vœux la création d'une agence numérique pour aller plus loin et valoriser au mieux ce nouvel atout", ajoute Renaud Fabre, plaidant pour la "nécessité d'un lieu de regroupement".



LE TDM : UN "TÉLESCOPE DE LA SCIENCE"



Concrètement, le TDM peut s'apparenter à un "téléscope de la science", décrit Claude Kirchner, permettant tout autant d'explorer la littérature scientifique publiée par les éditeurs scientifiques (il rappelle qu'un grand éditeur comme Elsevier publie entre 1,3 et 1,5 million d'articles chaque année), mais aussi les publications que les chercheurs accumulent dans leurs propres archives ouvertes. Et par "publication", il faut entendre non seulement les textes eux-mêmes, "mais aussi toutes les données d'expérimentation et d'usage", complète-t-il.

De son côté, Olivier Le Gall, directeur des affaires scientifiques de l'Inra, entrevoit que le TDM "amène aussi, in fine, à revoir le processus d'évaluation des chercheurs, en prenant en compte ce qu'apporte l'exploration du monde des données, en plus du monde réel. Le métier du chercheur change, il faut savoir le reconnaître et l'évaluer de façon différente", soutient-il. Claude Kirchner ajoute que "le TDM pousse davantage encore en avant la question de l'intégrité scientifique", dès lors qu'il faudra désormais s'attacher à "citer les données, les référencer, etc."

RÉINVENTER COMPLÈTEMENT LA QUESTION DE LA PUBLICATION

Enfin, Olivier Le Gall et Claude Kirchner pensent qu'il faut "réinventer complètement la question de la publication". "L'open access devrait être une nécessité", estime Olivier Le Gall. "Il y a des initiatives sur des systèmes de préprint, avec des réflexions sur ce qu'est la revue par les pairs. Il faudrait pouvoir publier directement en accès libre des textes qui seront directement exploitables par les machines. Cela va prendre quelques années, mais ça commence déjà en biologie ou en

SHS." "Il est fondamental de se réappropriier l'ensemble des connaissances que l'on produit, en sachant les documenter, les publier, etc.", renchérit Claude Kirchner.

En tant que chef de projet du projet Istex, associant "136 maisons d'édition", Renaud Fabre souligne avoir "à cœur que la qualité de la documentation scientifique devienne homogène" (lire sur AEF). "C'est un long processus, global et assez lourd", reconnaît-il. Il veut cependant croire à la "logique des communs", un concept développé par l'économiste Elinor Ostrom et pour lequel elle a reçu le prix Nobel d'économie en 2009 : "Plus on partage, plus on s'enrichit. Restons dans cette idée qui englobe toutes les logiques que l'on cherche à mettre en œuvre dans le TDM."

Exemples d'utilisation du TDM

Des microbes dans le fromage



Claide Nédellec, directrice de recherche à l'Inra, utilise le TDM pour chercher, dans la littérature scientifique, "les publications qui se sont déjà intéressées à telle ou telle bactérie dont on se demande comment elles se retrouvent dans un fromage". Le problème, sans l'autorisation donnée par la loi, c'est que cette fouille "extrêmement sophistiquée" n'est possible que sur "13 % des publications", faute d'avoir les droits de pratiquer le TDM sur les autres. Autre problème : "Je n'ai pas le droit de conserver ce que j'ai téléchargé." Outre la question de la reproductibilité, cette limite oblige aussi à repartir de zéro si d'aventure Claire Nédellec voulait réétudier le même corpus, dont la mise au propre pour le TDM a nécessité "un an de travail". "La loi est un premier pas formidable, mais un certain nombre de points restent à résoudre", conclut-elle.

Chercher de nouvelles espèces de poissons à domestiquer



Mathieu Andro, chargé de projets de numérisation et de text mining à l'Inra, a recours au TDM pour identifier des espèces de poissons sauvages que l'on pourrait espérer domestiquer. Il dit avoir pu négocier "rapidement" un contrat auprès de 24 éditeurs scientifiques, dont Elsevier, mais ce traitement éditeur par éditeur est "trop chronophage". Aussi préfère-t-il passer par les archives d'Istex, qui autorisent un traitement "plus simple" et permettent donc un "gain de temps". Selon lui, il est toutefois "possible que les documents fournis à Istex par les éditeurs ne soient pas d'aussi bonne qualité que ceux qu'ils diffusent sur leurs propres sites", ce qui rend le TDM plus complexe.