



Direction de l'Information
Scientifique et Technique

3 rue Michel-Ange
75016 PARIS

DIST : le fil 07-04-2015

Springer intègre à son workflow éditorial l'outil de détection de faux scientifiques SciDetect développé au Laboratoire d'Informatique de Grenoble

Springer Science+Business Media a annoncé le 23 mars avoir intégré dans son workflow éditorial l'outil *ScienceDetect*, développé par une équipe de chercheurs de Grenoble. Ce logiciel permet, en amont des tâches de *Peer Reviewing*, de déterminer si un article scientifique a été produit, tout ou partie, grâce à un automate de génération de texte. L'annonce de Springer est une importante reconnaissance pour les travaux menés au Laboratoire d'Informatique de Grenoble (LIG, www.liglab.fr), l'UMR 5217 associant l'Université Josef-Fourier, le CNRS et l'INP Grenoble. En février 2014 Springer découvrait que 18 articles publiés dans ses revues d'informatiques et d'ingénierie étaient des « faux » fabriqués par le logiciel *SciGen*, un générateur automatique de texte capable de produire des contenus ayant tous les attributs d'un article scientifique. Le N°2 mondial de l'édition scientifique a alors pris contact avec *Cyril Labbé* (1) un chercheur en informatique du LIG qui s'était taillé en 2010 une réputation personnelle en créant de toute pièce un chercheur-publiant virtuel baptisé *Ike Antkare* (comprendre : « *I can't care* ») auquel était attribuées une thèse (de 4,5 pages dont 2/3 de citations) et nombre de publications incohérentes issues d'un générateur de textes aléatoires aux apparences de travaux scientifiques. Le but était de mettre en évidence que les algorithmes de *Google Scholar* calculant des indices d'évaluation de chercheurs à partir de leurs publications laissaient à désirer : *Ike Antkare* en quelques mois se classait parmi les 100 meilleurs chercheurs au monde dans sa discipline.

Springer a décidé de financer les travaux d'un thésard de l'équipe de Cyril Labbé, Tien Nguyen, qui a développé *SciDetect*. La génération automatique de textes à partir de logiciels tel *SciGen*, un logiciel ouvert, mais aussi à partir de ses déclinaisons thématiques *Mathgen* ou *Physgen*, a multiplié les possibilités de « fabrication » de faux articles scientifiques. *SciDetect* est capable de repérer si un article scientifique est tout ou partie issu d'un générateur de texte. Springer va continuer de financer le développement de *SciDetect*, mais le logiciel est un logiciel libre, mis à disposition de tous sous licence ouverte GPLv3 (téléchargement : <http://scidetector.forge.imaq.fr/>).

(1) Lire : www.ujf-grenoble.fr/sites/default/files/KSUP/depeche33_1328523747888.pdf UMR 5217 –