

Séminaire Aristote Accréditation des résultats de la Recherche : La vérité est ailleurs?!

19 décembre 2014 École Polytechnique (Palaiseau)

Génération et détection automatique de faux articles scientifiques

Cyril Labbé (Université Joseph Fourier)

Il existe des générateurs aléatoires d'articles scientifiques, notamment en informatique mais aussi en physique, en mathématique ou encore en philosophie.

Ces articles, dépourvus de sens, peuvent être utilisés de différentes manières. Ainsi, ils ont permis à « Ike Antkare » de devenir l'un des scientifiques les plus cités, du moins si on en croit Google Scholar [1].

De tels articles apparaissent aussi dans des conférences réelles et par conséquent dans les services de bibliométrie les plus réputés (Scopus, ISI-Web of Knowledge) [2,3]. La présence de ces documents pose de manière aiguë la question de l'accréditation et/ou de la validation des documents et résultats qui sont publiés. On verra que ces générateurs aléatoires d'articles scientifiques sont assez faciles à identifier de manière automatique.

Bibliographie

[1] Cyril Labbé, "Ike Antkare, One of the Great Stars in the Scientific Firmament, 22th newsletter of the international society for scientometrics and informatics, June 2010.

[2] Cyril Labbé, Dominique Labbé. "Duplicate and fake publications in the scientific literature: how many SCIdgen papers in computer science?" *Scientometrics* 94, no. 1 (2013): 379-396

(<http://hal.archives-ouvertes.fr/docs/00/71/35/55/PDF/0-FakeDetectionSci-Perso.pdf>)

[3] Richard Van Noorden. Publishers withdraw more than 120 gibberish papers. *Nature*, February 2014.